

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2003, 19:40:10 ; Search time 35 Seconds  
(without alignments)  
2095.794 Million cell updates/sec

Title: US-09-675-509-4

Perfect score: 1802

Sequence: 1 MSQPKTLVGLFPLPSMN.....CVDPAKEVKNCAVLRPFL 356

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

1 number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL 21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	118.5	6.6	345	16	Q8RD78
2	116	6.4	1248	10	Q9SAJ2
3	115	6.4	286	2	Q9RLV6
4	114.5	6.4	1431	5	Q9XYH6
5	112.5	6.2	1360	5	015807
6	111.5	6.2	613	17	Q97YF0
7	111.5	6.2	1414	17	Q8TH79
8	111	6.2	417	16	Q8RBL7
9	110.5	6.1	718	10	Q94JF5
10	110.5	6.0	448	10	024399
11	109	6.0	998	5	044800
12	108.5	6.0	528	5	Q20730
13	108	5.9	2081	17	Q9HU37
14	107	5.9	1007	9	Q9FAJ3
15	106.5	5.9	1033	10	Q9CAA7
16	106.5	5.9	1033	10	Q9CAA7

17	106.5	5.9	1166	12	Q8V7K1
18	106.5	5.9	2819	16	Q980P8
19	106	5.9	545	17	Q9Y283
20	106	5.9	1054	16	Q9PR50
21	106	5.9	1220	3	Q9PJ38
22	105.5	5.9	358	16	Q928M1
23	105.5	5.9	602	16	Q8RBX6
24	105.5	5.9	1277	16	050698
25	105	5.8	1058	13	Q91943
26	104.5	5.8	603	17	Q97KR3
27	104.5	5.8	528	16	Q9PQ47
28	104.5	5.8	2462	16	Q8RG23
29	104.5	5.8	2806	16	Q8RI19
30	104	5.8	689	2	Q8RS20
31	103.5	5.7	692	17	Q97YM6
32	103.5	5.7	1733	17	Q8TTS9
33	103.5	5.7	4643	5	Q9VW71
34	103	5.7	286	16	Q9CIN5
35	103	5.7	637	16	Q92DM5
36	103	5.7	738	2	Q9AOF4
37	103	5.7	1857	3	Q8TCS5
38	103	5.7	2701	6	Q8W96
39	102.5	5.7	598	2	Q9L746
40	102.5	5.7	704	16	Q31399
41	102	5.7	578	10	Q9SDM7
42	101.5	5.6	505	5	Q93335
43	101.5	5.6	703	16	Q93MA0
44	101.5	5.6	728	10	Q9W9S4
45	101.5	5.6	767	5	Q94543

#### ALIGNMENTS

RESULT 1

Q8RD78 PRELIMINARY; PRT; 345 AA.

AC Q8RD78;  
 DT 01-JUN-2002 (TEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)  
 DE Sporulation protein and related proteins.  
 GN SPOL1 OR TBE0168.  
 OS Thermobacter tengcongensis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
 OC Thermobacteriales; Thermobacteriaceae; Thermobacter.  
 OC NCBI\_TaxID=119072;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MB4T / JCM11007;  
 RX MEDLINE=21992816; PubMed=11997336;  
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,  
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,  
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;  
 RT "A complete sequence of T. tengcongensis genome.";  
 RL Genome Res. 12:689-700(2002).  
 DR EMBL; AEO12990; AM23469.1; .  
 DR Complete proteome.  
 KW SEQUENCE 345 AA; B75DD6A0FB89AC57 CRC64;  
 SQ

Query Match 6.6%; Score 118.5; DB 16; Length 345;  
 Best local Similarity 21.3%; Pred. No. 3;  
 Matches 73; Conservative 60; Mismatches 118; Indels 91; Gaps 17;

QY 21 ENGSEVKILINIK-DVLEPTQVSGYN-IEYTEFDCYSDASLOS-LPDVFSSTDSI----- 70  
 DB 41 EGGKSVKDKLPSEYEVIRVFITNQNKIEEMQLEIDYKGVVAEMPAFEIEALKAQAVAA 100  
 QY 71 ---FLPVIYSVGKVSUDESIVRGVTDLHGFVSASSSVKSVGPGQYLCSNLFLLSPN 127  
 DB 101 RTYALAKEVALGGS-----KGC-DLHEGVD-----VCTD-----PE 129  
 QY 128 GTQOASSLLEIAKVG-----YEQI-----VYPDVASSSFTVFGUYOQLLO 169

Db 130 HQAQMOSBELKGMKNGENFEKYKISQAVDSTKGLVMVYQDA-----LILPVHAIISG 183  
Qy 170 SSSSAVDIKASDLPQSGDQVN--KDITQKRYRTILDSTVVASQREYINSVKQKPISNYY 227  
Db 184 GRTESEEDVMQCKIRPYLKSVSVPGEVASKYKT-----FTVVSQAEPVKLTKERQF--SLK 237  
Qy 228 VQSSSMCEIKIITIDQOYNVOLIGTSDKPY--VYTDVLALNS--NLCDERKVAVEYIK 283  
Db 238 LTSNNILBEIKIERTQAGHVTKLIGDVFEGKEIKELFGJNSTNFTFSKQK----- 290  
Qy 284 NLTLTIVLDLGLTLTPANKNGIAHLAKSSNPFYAOLSOOF 325  
Db 291 ----DDIVITVIGYGHGVMGQYGANALAKEGKFDILIKY 328

## RESULT 2

Q9SAJ2 PRELIMINARY; PRT; 1248 AA.

Q9SAJ2  
ID 01-MAY-2000 (TremBrel. 13, Created)  
DT 01-MAY-2000 (TremBrel. 13, Last sequence update)  
DE 01-MAR-2002 (TremBrel. 20, Last annotation update)  
DE TRK14.1 protein.  
GN TRK14.1.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV\_COLUMBIA;  
RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,  
RA Li J., Kremenetskaia I., Lueros J., Lee J.M., Gonzalez A., Altai H.,  
RA Araujo R., Chao Q., Conn L., Conway A.B., Dunn P., Hansen N.,  
RA Huizar L., Kim C., Palm C., Rowley D., Shinn P., Walker M.,  
RA Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;  
RT "Arabidopsis thaliana chromosome 1 BAC TRK14 sequence."  
RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AC007202; AAD30219.1; -.  
DR HSSP; P08631; 1AD5.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR004040; STY\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR PRODOM; PD000001; Euk\_pkinase; 1.  
FT SMART; SM00221; STYK; 1.  
LE PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS0011; PROTEIN\_KINASE\_DOM; 1.  
KW ATP-binding; Transferase; Tyrosine-protein kinase.  
SQ SEQUENCE 1248 AA; 137242 MW; 0EE4A8B99C50328 CRC64;

Query Match 6.4%; Score 116; DB 10; Length 1248;

Best Local Similarity 19.1%; Pred. No. 27;  
Matches 75; Conservative 69; Mismatches 123; Indels 126; Gaps 17;

Qy 14 FYLPSW-----NENGNEVKLN-----LIKD-VLPTQVS----- 41  
Db 26 PFMPDQTSASANMRPPNSNGSDVKAHNFISIQTEEFSLERMRDVIIPQRSSNPNGAD 85  
Qy 42 -GYNIEYEF-----DCVSDASLQS-----LDVFTDLSIFLPYVLSGCVKSL 84  
Db 86 MMYNTGYELKGLIGISHTGSECADVSRFSTVENGTSDIERTNSLHEFGKLNHVQSA 145  
Qy 85 DESIVR--GVGTGLHSFVSSASVNGSVYGFPOYLCS--NFLSSPNNGTQ-----QASGL 135  
Db 146 FOALLSKSSSGNLHGYNKNTSSASGSVTAKYLCSFGKILPRPGSKLRYVGGEFTHI 205  
Qy 136 LELAKVGYEQT-----VYPDVASSSFTVFGLYQ-----QLIQSSSSAAVDIKASD 182  
Db 206 ISIRKDISMOELROKILRY-----YQTRVVKYQLDGEDLDLALVSVSSSED 251

Qy 183 ----LPSGDQVNDITQKRYRTILDSTVVASQREYINSVKQKPISNYYGVSMSCEIK 238  
Db 252 LQNNLEEVNENENGGSKQKMLFVS--ISDMDALLGVNNDSDSEF----- 297  
Qy 239 DIIRHQOYNVOLIGSDKPYVYTDVLALNSNLCDERKVAVEYIKNLTLTIVLDLGLTL 298  
Db 298 -----QTVVAVNGMDISGKNSITLIGDSSANNLAELDVNRNTEGI--NTIAGDVGVG 349  
Qy 299 LTLPAKNGIAHLAKSSNPFYAOLSOQFPDAKES 331  
Db 350 -----ASQLMVNGFOQTSAQOSE 367

## RESULT 3

Q9RLV6 PRELIMINARY; PRT; 286 AA.

Q9RLV6  
ID 01-MAY-2000 (TremBrel. 13, Created)  
DT 01-MAY-2000 (TremBrel. 13, Last sequence update)  
DE 01-JUN-2002 (TremBrel. 21, Last annotation update)  
DE Hypothetical 32.0 kDa protein.  
OS Lactococcus lactis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
OC Streptococcaceae; Lactococcus.  
OX NCBI\_TaxID=1358;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MG1363;  
RX MEDLINE=20112344; Pubmed=10647818;  
RA Aungmyaphanoponchai P., Griffin H.G., Gasson M.J.;  
RT "Cloning, DNA sequence analysis, and deletion of a gene encoding  
RT diacetyl-acetoin reductase from Lactococcus lactis."  
RL DNA Seq. 10:163-172 (1999).  
DR EMBL; AJ012388; CAB59827.1; -.  
DR InterPro; IPR004872; Lipoprotein\_9.  
DR Pfam; PF03180; Lipoprotein\_9; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 286 AA; 31960 MW; 4B481A83D4314B CRC64;

Query Match 6.4%; Score 115; DB 2; Length 286;

Best Local Similarity 22.2%; Pred. No. 3.9;  
Matches 47; Conservative 42; Mismatches 85; Indels 38; Gaps 7;

Qy 13 FPLYPSWNE--NENGNEVKLNIKVLPVQVSGNIEYTFPCYDASLQSLPDPFSTDSI 70  
Db 99 YAYLKSNNKANNNGNIVS---IGDTITTPHLVSTKYKVDLPDKSTIAIPNDITNESR 154  
Qy 71 FLPLVLSGCVKSDSEIVRGVGTGLHSFVSSASVNGSVYGFPOYLCSNFLSSPNNGTQ 130  
Db 155 AL-VYLNKAGLIKLDTS--KGLATVADIRENPKSL-----IIEIDASQ 196  
Qy 131 QASLLEIAOKRYGEQIVPDVASSSFTVFGLYQQLQSSSSAAVDIKASDLPQSGDQV 190  
Db 197 TPRALDSVAAV---INYNPAISAKNSDKESIVQEPLNDSAQWIFIANQSDKNKKV 252  
Qy 191 NKDITQKRYRTILDSTVVASQREYINSVKQKRP 222  
Db 253 YKEVAVK-----EQKNIAIDITKEYP 274

## RESULT 4

Q9XYH6 PRELIMINARY; PRT; 1431 AA.

Q9XYH6  
ID 01-NOV-1999 (TremBrel. 12, Created)  
DT 01-NOV-1999 (TremBrel. 12, Last sequence update)  
DE 01-JUN-2002 (TremBrel. 21, Last annotation update)  
DE ATP-binding cassette protein.  
GN CPABC.  
OS Cryptosporidium parvum.  
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;  
OC Cryptosporidiidae; Cryptosporidium.

Query Match	6.4%	Score 114.5;	DB 5;	Length 1431;
Best Local Similarity	22.9%	Pred. No. 42;		
Matches 89; Conservative	61;	Mismatches 142;	Indels 97;	Gaps 21;

RESULT 5	
015807	
ID 015807	PRELIMINARY;
	PRT; 1360 AA

DT 01-JAN-1998 (TrEMBLrel.05, Created)  
 DT 01-MAR-2002 (TrEMBLrel.20, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel.21, Last annotation update)  
 DE ATP-binding cassette protein 1 (Fragment).  
 GN ABC1.  
 OS Cryptosporidium parvum.  
 OC Eukaryote, Alveolata, Apicomplexa, Coccidia, Eimeriida  
 OC Cryptosporidiidae, Cryptosporidium.  
 NCBI\_TaxID=5807;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=KSU-1;  
 EX MEDLINE=97376910; PubMed=92333681;

Query Match	6.2%;	Score 112.5;	DB 5;	Length 1360;
Best Local Similarity	22.9%;	Pred. No. 53;		
Matches 89;	Conservative 61;	Mismatches 142;	Indels 97;	Gaps 21

RESULT 6  
Q97YF0  
ID Q97YF0 PRELIMINARY; PRT; 613 AA

DT 01-OCT-2001 (TEMBLRef1 . 18, Created)  
DT 01-OCT-2001 (TEMBLRef1 . 18, Last sequence update)  
DT 01-OCT-2001 (TEMBLRef1 . 18, Last annotation update)  
DE Hypothetical protein SSO1375.  
GN SSO1375.  
OS Sulfolobus solfataricus.  
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae  
OC Sulfolobus.

OX NCB1\_TaxID=2287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 35092 / DSM 1617 / P2;  
RX MEDLINE=21332296; PubMed=11427726;  
RA She O., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,  
Ra Aweyer M.J., Chan-Weher C.C.Y., Clausen I.G., Curtis B.A.,  
Ra De Moors A., Erasus G., Fletcher C., Gordon P.M.K.,  
Ra Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,  
Ra Thi-Noc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,  
Ra Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,  
Ra Garrett R.A., Regan M.A., Sensen C.W., Van der Oost J.,  
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).  
DR EMBL; AE006751; AAK1609.1;  
KW Hypothetical protein; Complete proteome.  
SO SEQUENCE 613 AA; 67767 MW; F478640AC69BA3 CRC64;

Query Match 6.2%; Score 111.5; DB 17; Length 613;  
Best Local Similarity 19.5%; Pred. No. 20;  
Matches 85; Conservative 72; Mismatches 113; Indels 165; Gaps 22;

QY 13 FPY---LPSMNGN-----NEVKLIN-ISKVLPYVSGYV-----44  
DB 88 YPIQVWNYNPNGLPLMNSYDNMSIHESIFALPTVDVYPTKI--YNSSLYLKGLLS 145  
QY 45 -----LEYTEFDCYSASLQSLPD--VFSTDSIFLPYVLSLGV---KSLDES 87  
DB 146 ATYKNIATIIISYKDFSKY-NITLSQYPSHILQGTGIGL-PIIPLOALPMWYKGSVNS 203  
QY 88 LVR-----GVTDLHSFVSSASV-N-GSYVGFPOYVCSNPL 122  
DB 204 VIQGYIFQNETNLINGNDPGLTOLGLYDELSTVQTLFELNTSPY-----SSIL 256  
QY 123 LSPSPGTOA-----SSLELA-----OKVGEQIV-----148  
DB 257 IDSINNTLEAANRTTINOTINDSLIDAYSSSTLTPEGEFFAYIVSPGNITTKIES 316  
QY 149 -----YPDVASSSSFTFGIYQOOLQSSSSAAVDIKASDLPQSGDQVNDKI 194  
DB 317 EYISLPKLVMLIQYPSISNOSPITIVSFVFNLRSAATKVEGVYGLIYNSG-----369  
QY 195 TOKRTIIDSTVVASQREYINSVKQKPISNY-----VGYSEMCIEKDIIRD--QOY 246  
DB 370 -----LISKSIKMSNGELVFNVSPTVLYVYHYNLGLNTEYWGQIRIDVRSKINKY 423  
QY 247 NVQIGTSDKPYVYTDVLAUNSLCDEKQKAVEYIKNLTNTVLVDLGLGLTPANKN 306  
DB 424 NF---TVEEPIYNIIT---DTSIDQPKISVKIINPL--NOTVSGQLYLWIS-----N 468  
QY 307 GIAHLAKSSNFYAOL 321  
DB 469 GIASASAPTFESOL 483

RESULT 7  
Q8RH79 PRELIMINARY; PRT; 1414 AA.  
AC Q8RH79;  
DT 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Hypothetical protein MA4643.  
GN MA4643.  
OS Methanosarcina acetivorans.  
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;  
OC Methanosarcinaceae; Methanosarcina.  
OX NCB1\_TaxID=2214;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C2A / ATCC 35395 / DSM 2834;  
RX MEDLINE=21929760; PubMed=11932238;  
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,

RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Attoor D., Brown A.,  
RA Allen N., Naylor J., Scange-Thomann N., DeAtelliano K., Johnson R.,  
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,  
RA Limer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,  
RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,  
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,  
RA Springer T.A., Umayam L.A., White O., White R.H., de Maccario E.C.,  
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,  
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,  
RA Metcalf W.W., Birren B.,  
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic  
and physiological diversity";  
RL Genome Res. 12:532-542(2002).  
DR EMBL; AB011187; AAM07977.1;  
KW Hypothetical protein; Complete proteome.  
SO SEQUENCE 1414 AA; 159146 MW; D6F0D6CEB589A89 CRC64;

Query Match 6.2%; Score 111.5; DB 17; Length 1414;  
Best Local Similarity 21.7%; Pred. No. 65;  
Matches 83; Conservative 56; Mismatches 148; Indels 95; Gaps 19;

QY 15 YLPSPMNGN-----EVKLINLIKDVLPYVSG---YNIETEFDCYSASLQSLPD 63  
DB 649 YLTS-KENNNTALMDQYENISIRIKDTHIDEDIGLNFYSNMSYEFE---DVIYGEVH 704  
QY 64 VFSTDSIF--LPYVLSLGVKSLDES--LVRYTGDLHSFVSSASVNGSVYGFPOYLC 118  
DB 705 DYLDHMYANIPYGRIFAPRIPISDESAYALVRGMLGQ--SYIEHIQAVNAS-----C 754  
QY 119 SNFLSSPNCQOQASL-----LEIAQKVEYEDIVPVDVSSSSFTVFG--LYQOL 167  
DB 755 DSHETEYKNTESRYRLTYLVVDGIDITLLAO---BEVLQSSSAEVEDLVAEYYQNI 810  
QY 168 LQSSSSA-----AVDIKASDLPQSGDQVND--ITQKRTIIDSTVVASQREYINSVK 218  
DB 811 LLSSESEIKGLDLDAKYIEPSVGDILSPVLPKPKYSPDRPLIPKEXY-NIGK 869  
QY 219 QGKPISNYVYGVSSSMCEIKDIIRDQYVQVQLIGTSKPY-----258  
DB 870 E--TIDSFVLDVYKNEPEPEKVAVLMGIEYMRNHGVPHSQMLYLVGVEPWDMANGRTV 927  
QY 259 VYTDVLAUNSLCDEKQKAVEYIKNLTNTVLVDLGLGLTPANKNGIAHLAKSSNFY 318  
DB 928 YTTPSASNLHINQSE-LTIRLSNGTVINRPRIIDVIG-----HSSGLH 970  
QY 319 A-QLSQOFDAKSESEVRLRQVD 339  
DB 971 RDQFPQMDLLDPAVRIISQLD 992

RESULT 8  
Q8RH79 PRELIMINARY; PRT; 417 AA.  
AC Q8RH79;  
DT 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Sugar-binding periplasmic proteins/domains.  
GN UGPB OR TTE0799.  
OS Thermobacterium tengcongensis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
OC Thermobacteriales; Thermobacteriaceae; Thermobacter.  
OX NCB1\_TaxID=119072;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MB4T / JCM11007;  
RX MEDLINE=21992816; PubMed=11997336;  
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,  
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,  
RA Tan H., Chen R., Wang J., Yu J., Yang H.,  
RT "A complete sequence of T. tengcongensis genome";  
RL Genome Res. 12:689-700(2002).  
DR EMBL; AE013046; AAM24056.1; -.

KW Complete proteome.  
SQ SEQUENCE 417 AA; 46304 MW; C974291CB505202D CRC64;  
Query Match 6.2%; Score 111; DB 16; Length 417;  
Best Local Similarity 20.8%; Pred. No. 13;  
Matches 83; Conservative 68; Mismatches 143; Indels 106; Gaps 19;  
QY 2 STOPKTLVGLPFPVLPWMNNGNEVKLINLIKDVLPQTQVSGNIEYTEFD----- 51  
DB 38 TSKPTVTLGLM-----WSSSPAEEKIYD-----DGIARFKKYPRIIDVQIETIVGDY 84  
QY 52 ---CYSDASLQSLDPVSTDSIFLPYLVSLGVKSLDLSLVGVGTGLHSF---VSSSAS 105  
DB 85 MQKLQTEIASNTAPDIFYLDSDMPAPQLMSSGVLEPLDE-YIKKYVVDVNDPEPALLSAFQ 143  
QY 106 VNGSVYGFPOYLCSNFISSPNGTQA-----SSLLELAQV---GYEQIYY-PDV 152  
DB 144 WEGKTYGLPKDYNLTVLFPYNDMPKEAGINPPKTMWEELRETAKKLTQKGVGLSADL 203  
QY 153 ASSSFTVFGLYQQLQSSSAVDIKAS-DLPQSGDQV-----KDITQKXRT 200  
DB 204 ARPDFA-----INQDGSVYKDGKVTLLPQNAELDPVGLIIRDKVADTFQNNGE 255  
QY 201 ILDSTVASQRE-----YINSVKQKPISNYVYVSGESMCEIKDIRDOQYVWQLIG 252  
DB 256 GMDGDAFAKKAAMIEGGMWIPFLKEKAPDLNY-GIAELPAKCK----- 299  
QY 253 TSDRPYVTDVLALNSLCEKQKAVAVIKLTLNTLVLDLGLGLPANKGIAHLA 312  
DB 300 KSTYAFVAYVMNKNKSKHDEAFKL-IEFLTGKEGQGVVD---SGLALPSRK-----S 349  
QY 313 KSNF---YAQLSQQFPAKESEV-----RVLRQVDPFANK 343  
DB 350 MEANFKKYPERRAPRIDASVAVPQFGLYGTKLVDAAANK 389  
RESULT 9  
QY 094JF5 PRELIMINARY; PRT; 718 AA.  
AC 094JF5;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE P0684B02.13 protein.  
GN P0684B02.13.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Euphorbiaceae; Oryzaeae; Oryza.  
NCBI\_TaxId=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=CV, NIPONBARE;  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa niponbare(GA3) genomic DNA, chromosome 1, PAC  
clone:p0684B02.13";  
RU Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF003023; BAB44026.1;  
DR InterPro; IPR000644; CBS\_domain.  
DR InterPro; IPR001807; CI-channel\_volt.  
DR Pfam; PF00571; CBS; 2.  
DR Pfam; PF00654; voltage\_CLC; 1.  
SQ SEQUENCE 718 AA; 76362 MW; DA854F202012785B CRC64;  
Query Match 6.1%; Score 110.5; DB 10; Length 718;  
Best Local Similarity 26.2%; Pred. No. 29;  
Matches 71; Conservative 36; Mismatches 93; Indels 71; Gaps 15;  
QY 55 DASIQSLDPVSTDSIFLPYLVSLGVKSLDLSLV--GVTDLHSFVSSASVNGSVYG 112  
DB 367 DILLESRP---FTSGLSAAVIVQILGVKVLATSLCRAFGLGVGYAPSLFICAAITGMAYG 423  
QY 113 PPOYLCSNFISSPNGTQAASLLELAQKVGVEQIVYPDVASSSSFTVFGLYQQLQSSS 172

DB 424 ---KVMRFPTGPD-----SLF-----QIPFLDVASPOAYGLVGM----- 455  
QY 173 SAADVIRASDLPOSGDQVNDKITQKRTIILD-----STVASQREVINVKQKQPISN 225  
DB 456 -AATLAGVCKPPLTSVLLFELTDYRIVLPGLGAVGVSWIASPQPSRS-NSKSPSS 513  
QY 226 YVGVSESMCEIKDIRDOQYV--QLIGTS---DKPYVTDVLALNSLNC--DEKOK-- 276  
DB 514 -----EVKSNRQKQESVPSQTOGASVDIDKP--ITDLCKLSSLCVPAKHENF 561  
QY 277 ---VAVEVIR---NLTLNTVLDDLGLGL 299  
DB 562 QENTLVAEMKTKYISVSKTPVVEALNLM 592  
RESULT 10  
QY 070410 PRELIMINARY; PRT; 912 AA.  
ID 070410;  
AC 070410;  
DT 01-AUG-1998 (TREMBLrel. 07, Created)  
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Putative pheromone receptor V2R2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxId=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=VOMERONASAL NEURONS;  
RX MEDLINE=97436753; Pubmed=9292726;  
RA Ryba N.J., Tirindelli R.;  
RT "A new multigene family of putative pheromone receptors.";  
RN Neuron 19:371-379(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=VOMERONASAL NEURONS;  
RA Ryba N.J.P., Tirindelli R.;  
RU Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF053986; AAC08413.1;  
DR InterPro; IPR001828; ANF\_receptor.  
DR InterPro; IPR000337; GPCR\_Mgr.  
DR Pfam; PR00003; 7tm\_3; 1.  
DR Pfam; PF01094; ANF\_receptor; 1.  
DR PRINTS; PR00248; GPCRMR.  
DR PROSITE; PSS0259; G\_PROTEIN\_RECP\_F3\_4; 1.  
KM Receptor.  
SQ SEQUENCE 912 AA; 102348 MW; 2C54FAB6DBFA48D CRC64;  
Query Match 6.1%; Score 110.5; DB 11; Length 912;  
Best Local Similarity 21.0%; Pred. No. 41;  
Matches 79; Conservative 62; Mismatches 120; Indels 115; Gaps 20;  
QY 26 VKLINLKDVLPQTQVSGNIEYTEFD-CYS-----DASLQSL-----PDVFTDSIFL 72  
DB 95 IKETIKRDLIPNHTLTYQI---FDSCYITISKAMESSLVFLTGQEEFKPFRNRSTGSL 150  
QY 73 PYLVSLGVK-SLDESIVRGVGTGLHSFVSSASVNGSVGFPQYLCSNFISSPNGTQ 131  
DB 151 AALVSGSGSSLSVAASRIILGLYVWPQVYTSSCGILSDKQFPYS-----LRVPSDNLQ 205  
QY 132 ASSLLELAQKVGVEQIVYPDVASSSFTVFG--YQQLQSS-----SSAAVDIKASDL 183  
DB 206 SEAVINLIKFGW--VWVGAIADDDYGVKGVKFKKEMESANLVAFSERIPRVYSEK 263  
QY 184 PQSDQVNDKITQK-----YRTLDSTVASQREYINSVKQKPI 224  
DB 264 MQKAVKAVKSTAVIVLYTSDIDLSFLVEMIHNTIDRTWIAE-AWITSALIAKP-- 320  
QY 225 NYV-----GVSESMCEI---KDIIHQDQ-----YN 247  
DB 321 EYFYPGQTIGFAIPRSVITGLKFLYDVHPNDPNDVLTIEFWQTAFNCTWPNSSVPYN 380

QY 248 ----VQIGTSPKPYVYTDVLAALNSLCEKQKVAVEVIAKLLTNTVLDLGLGLT---- 300  
 Db 381 VDHRRNMGMKEDERLYDMSD-----QLCTGEK--LELDKN-----TYLDTSQRLRTKQC 427  
 QY 301 ---LPANKNGIAHLAK 313  
 Db 428 KQAVYAIHAGLDHLSR 443

RESULT 11

024399 PRELIMINARY; PRT; 448 AA.  
 ID 024399  
 AC 024399  
 DT 01-JUN-1998 (TREMBlrel. 05, Created)  
 DT 01-JUN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Fimbrin/plastin-like (Fragment).  
 OS Triticum aestivum (wheat).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 OC Triticeae; Triticum.  
 OX NCBI\_TaxID=4565;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV, VICTORY, TISSUE=ROOT;  
 RX MEDLINE=97422889; PubMed=9276954;  
 RA Cruz-Ortega R., Cushman J.C., Ombry J.D.;  
 RT "cDNA clones encoding 1,3-beta-glucanase and a fimbriin-like  
 cytoskeletal protein are induced by Al toxicity in wheat roots.";  
 RL Plant Physiol. 114:1453-1460(1997).  
 DR EMBL; U67717; AAC49613.1; -.  
 DR HSSP; P13797; IAOA.  
 DR InterPro; IPR001589; Actbind\_actnin.  
 DR InterPro; IPR001715; Calponin-like.  
 DR Pfam; PF00307; CH; 4.  
 DR SMART; SM00033; CH; 3.  
 DR PROSITE; PS00019; ACTININ\_1; UNKNOWN\_1.  
 DR PROSITE; PS50021; CH; 3.  
 FT NON TER 1  
 SQ SEQUENCE 448 AA; 50505 MW; 3C0A39B8DA808C7 CRC64;

Query Match 6.0%; Score 109; DB 10; Length 448;  
 Best Local Similarity 20.7%; Pred. No. 19;  
 Matches 77; Conservative 72; Mismatches 155; Indels 68; Gaps 17;

QY 16 LPSNNGNGEVLINLIDV-----LPTQVSGYNIEYEPDQCYSDA-----SLQ 60  
 F 5 LNPFRNENHNLCLNSAALIGCTVNTIGTQFLVEGRPHLVGLISQITIKIQLLADLNLKK 64  
 QY 61 LPDVS--TDSIFLPLVLSLGVKSLDESILVR-----GVTDLHSFVSSASVNGS 109  
 Db 65 TPQLVELPDDSKDIDEVLSL-----STERKMLRWMMHHLKKAGYKKTANNF--SSDYKGE 118  
 QY 110 VYGF-----PQYLCNPLSSPNTQOASSLLELAQKVGEQIYYP-DVASSSFTYVFG 162  
 Db 119 AYAALLKALAE-TSPETTLERKNPDERAKWLEQAERKIDCRKRYLTPDIDTEGSANLMLA 177  
 QY 163 LYQQLSSSSAAVDIKASDLPOSGDQVKNKIDTK---YRTILDSTVVASQREYINSYKQ 219  
 Db 178 FVAQIFQRNGITSDIKQVTLTQASRPDVLVSREERAFRMKNTSLGVS--YNNVFE 234  
 QY 220 GKPSINYYGVSESMCEIKDIIIRDQYVVOVLIGTSDEKPYVYTDVLAALNSLCEKQKVAV 279  
 Db 235 --DVNRNGV-----LLEVLDKVSPSSVWKL---ESKRPILKPFKLEN--CNQVYKIGK 282  
 QY 280 EV---IKVLLNTLVL--DLGLGLTPANKNGIAHLAKSSNFAQLSQQPDASEVAV 334  
 Db 283 ELKESLVMLAGNDIVQGNKKLVALLMQIMRFNLIQLNRLRSHSGSGQKQITADI-- 340  
 QY 335 LRCVDPAKKEVY 346  
 Db 341 ---LNMANSKVK 349

RESULT 12

044800 PRELIMINARY; PRT; 998 AA.  
 ID 044800  
 AC 044800  
 DT 01-JUN-1998 (TREMBlrel. 06, Created)  
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE F14D2.6 protein.  
 GN F14D2.6.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
 OC Rhabditidae; Peloiderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copesey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Larellie P.,  
 RA Latching J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
 RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulton J.,  
 RA Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkison-Sproat J., Woldman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans.";  
 RL Nature 368:32-38(1994).  
 DR EMBL; U67717; AAC49613.1; -.  
 DR HSSP; P13797; IAOA.  
 DR InterPro; IPR001589; Actbind\_actnin.  
 DR InterPro; IPR001715; Calponin-like.  
 DR Pfam; PF00307; CH; 4.  
 DR SMART; SM00033; CH; 3.  
 DR PROSITE; PS00019; ACTININ\_1; UNKNOWN\_1.  
 DR PROSITE; PS50021; CH; 3.  
 FT NON TER 1  
 SQ SEQUENCE 998 AA; 113094 MW; BF48A4E614F90D1B CRC64;

Query Match 6.0%; Score 108.5; DB 5; Length 998;  
 Best Local Similarity 22.6%; Pred. No. 64;  
 Matches 76; Conservative 56; Mismatches 125; Indels 79; Gaps 17;

QY 8 LTVGLFPLPSNNGEVLINLIK-DV-LPTQVSGYNIEYEPDQCYSDA-----SLQ 59  
 Db 137 LITISYNSIDWWSQETSKWIDINNTKLDVGICVCAARDI-YADFNAYGNKKVVAADSYK 195  
 QY 60 SLDPVFETDSIFLPLVLSLGVKSLDESILVRGVTG-DLHSFVSSASVNGSVYFPQ-YL 117  
 Db 196 STTQGFETIQALQYLSLSLTKKLNGLMEIYETGLENVFLEGLLETHNGNGSLPKEYW 255  
 QY 118 CSNF--LSSPNTQOASSLLELAQKVGEQIYYPDVASSSFTV-----FGLYQOL 167  
 Db 256 SFNFTSIHDNN-----LRRGLDLSL--KQGPAGKFTIIRANNHPDFCLSTSE 302  
 QY 168 LQSSSSAAVDIKASDLPOSGDQVKNKIDTK---YRTILDSTVVAS 209  
 Db 303 LQVAFRTNLKIYGEV-----QICKDLFRDQGTCTCFKLSMLDPKCOHIIQIFINAT 357  
 QY 210 QREYINSVKQKPLSNV--YVGYSESMCEIKD-----IIRDQ---YNYQLIGTSPDKY 258  
 Db 358 NEETILNKKNTKYICGFFGAGFAGASCFYGEERIGDYLILFPSSKLLRNVTLLPQMKPPF 417  
 QY 259 VYTD--VLAALNSN-----LCDEKQKVAVEVIAKLLTNTVLDLGLGLT---- 300

Db 418 SPTNPNVIMNNNTGEIQTREACQEFQKFTKSIK 453

RESULT 13

Q20730  
ID Q20730 PRELIMINARY; PRT; 528 AA.  
AC Q20730;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE F53F4.5 protein.  
GN F53F4.5.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
OC Rhabditidae; Petriodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Barlow K.;  
RN [2]  
R Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; Pubmed=9851916;  
RA none;  
RT "genome sequence of the nematode C. elegans: A platform for  
investigating biology."  
RL Science 282:2012-2018(1998).  
DR EMBL; Z77663; CAB01214.1; -  
DR InterPro; IPR000960; Flav\_cont\_mnoxn.  
DR Pfam; PF00743; FMO-1like; 1.  
SQ SEQUENCE 528 AA; 60119 MW; C6261F9BFE37CE20 CRC64;

Query Match 6.0%; Score 108; DB 5; Length 528;  
Best Local Similarity 21.7%; Pred. No. 28;  
Matches 53; Conservative 47; Mismatches 80; Indels 86; Gaps 15;

QY 43 YNIEYEDCYSDASLSPVFTSDSIFLYLV-----SLGVK----- 82  
Db 238 YDVGQ--FSRYVDLTKRIPHAVAND--FMERYLQGRMDHYGKRPFRFOQHPTND 293  
QY 83 SLDSLVRG--VTGDLHSFVSSASVNGS-----VYGPQYLCNFISS 125  
Db 294 ALANLTCAGYITTEDIDTFENSIVYVKGREPKDIFLTCTGYTFGFP-----FV--- 344  
QY 126 PNGQAQASLLELA-QKVGQIYVPDVASSSFTVFGYQL-----LQSSSAAY 176  
Db 345 -----DSDIVEIKNQVPLKYVFP--PNSDVAIVGLIOPISIAISEIQRMARV 396  
QY 177 DIKASDLPOSGDQVKNKDTQKVRTLDSTVVASQREYINSYKQGPISNYVGYSESWC 236  
Db 397 FAGCGQPSSEGOID-DIQRK-----KAMMKRYRDSIKH--TIQVDYMSYDEIAE 445  
QY 237 IKDIIRDOQV-----VOLIGTSKPYVY 260  
Db 446 IICGLPMKHVLFYPRFPMKLFMGANVPYAY 477

RESULT 14

Q9HJ37  
ID Q9HJ37 PRELIMINARY; PRT; 2081 AA.  
AC Q9HJ37;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
DE Conserved hypothetical membrane protein.  
GN TAIL36.  
OS Thermoplasma acidophilum.  
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;  
OC Thermoplasmataceae; Thermoplasma.  
OX NCBI\_TaxID=2303;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=DSM 1728;  
RX MEDLINE=20479972; Pubmed=11029001;  
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,  
RA Mewes H.-W., Fishman D., Stocker S., Lupas A.N., Baumeister W.;  
RT "The genome sequence of the thermophilic scavenger Thermoplasma  
acidophilum."  
RL Nature 407:508-513(2000).  
DR EMBL; AL445066; CAC12262.1; -  
KM Complete proteome.  
SQ SEQUENCE 2081 AA; 227495 MW; 6E02AA6470DB2E8D CRC64;

Query Match 5.9%; Score 107; DB 17; Length 2081;  
Best Local Similarity 19.8%; Pred. No. 2,3e+02;  
Matches 77; Conservative 59; Mismatches 148; Indels 104; Gaps 15;

QY 15 YLPSMNENGVKILNLIKDVLPYQVSGNIEYTFPCYSDASLQSLPVSFSDSIFLP- 73  
Db 1102 YVP-----MNTINLALSLQGVNLGFSI-YVNGTAYSNSTYKN-----GHYLYLPS 1147  
QY 74 --YLVSLGVKSLDESIVRGVGTGLHSFVSSASVNGSVYGFPOYLCNFISSPNTQQ 131  
Db 1148 GNYTISARSSNATTDSDIYLPQ--WNSFVAVSLPYKAV-----RVSGTAS 1192  
QY 132 ASSLLELAQKVGVEQI-----VYDVASSSFTVFGI--YQQLQSSSAVADIK 179  
Db 1193 NISWIKFLSGNGVNVAVSVNTTGVPSVIVPAGTVYVYGQGVAFKTYTLTSNTDIN 1252  
QY 180 ASDLPQSGDDQVNDIT-----VYDVASSSFTVFGI--YQQLQSSSAVADIK 206  
Db 1253 ISGVPAATVSLSNITNIVSSGYSIVSSGSFYTYTSSGSQFSLSPSGYTYTYSAST 1312  
QY 207 VAS-----QREYINSVKQKPISNYVGYG--ESMCEIKD-IIRDOQV 247  
Db 1313 IAGSYSSILKPLTKTSVSYQMLHVNSTMGFVLYNAGISSVSNVSDGYIMLYSYG 1372  
QY 248 VOLIGTSKPYVYTDVLAHNSNLCDERQKVAVEVIXNLINTLVLLGLGLTPANKG 307  
Db 1373 IPVAVPINSQGYDVIYAPNIGSPKIEIISPDYNTLTN--ISSAMTIGLT-PMNVVP 1429  
QY 308 IAHAKSSNF-----YQQLQSPDAKES 330  
Db 1430 TITLYNSSRLAVFSGVAVLQGTVDYKLS 1457

RESULT 15

Q9F4J3  
ID Q9F4J3 PRELIMINARY; PRT; 1007 AA.  
AC Q9F4J3;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE PBLA.  
GN PBLA.  
OS Streptococcus mitis phage SM1.  
OC Viruses.  
OX NCBI\_TaxID=157924;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21116945; Pubmed=11179301;  
RA Bensing B.A., Rubens C.E., Sullam P.M.;  
RT "Genetic loci of Streptococcus mitis that mediate binding to human  
platelets."  
RL Infect. Immun. 69:1373-1380(2001).  
DR EMBL; AY007505; AAC18638.1; -  
SQ SEQUENCE 1007 AA; 107445 MW; E568CF7E242A792 CRC64;

Query Match 5.9%; Score 106.5; DB 9; Length 1007;  
Best Local Similarity 21.3%; Pred. No. 88;  
Matches 78; Conservative 53; Mismatches 114; Indels 121; Gaps 20;

QY 90 RGVTDGLHSFVSSASVNGSVYGFPOYLCNFI-----LSPNGTQ 130  
Db 17 RGITGKIQSIINPEASAAQAGSAG--QSLGSSLVGVMTKVIYAAGIGKAFSAISGALQ 74

```

QY 131 QASSLLEL-----AQV-GYEQIVYPDVASSSSF---TVFGLYOQLLOS---SSSAVVDI 178
Db 75 QSLGIEFTLFRKSADKVKYGVANBAVKTTGLSANAYMENVTGFSASLLQSLGDTNKAAET 134
QY 179 KASDLPOSGDQVNDITQKRTILDSTVVASQREYINSVKOG-KPISNYVGYSESMCEI 237
Db 135 ANMAMIDMSDNANK-----MGTSMESIOMAYQGFQKQNTMLDNLKLGYGGTQDEM 185
QY 238 KDIIQDOO-----YVQLIQSTDPKYVYTDVLALNSNL-----CDEKQ 275
Db 186 QRLADAEKLTGVKYDIN--NLSD---YSAIHAIQENLDITGTTAKAEAATFGSGFESM 240
QY 276 KVAVEVI-----KNLL-----TNTLVLD-----LLGLGLTPANKNGIAH 310
Db 241 KAAQNVVLKGLALGENILPSLHALLKTSTFLFDNFLPMIGNVFGSLGLVL--TEGISQ 297
QY 311 LAK-----SSNRYAQSQ-----QFD-----AKESEVRVLRCDVPFANK--EYKN 347
Db 298 IASQLFGDAPGSAVFDQLSRITGIFETFFDMLFGSLSKQDNIDILNTIGFSEEAATQIVN 357
QY 348 CAGVLR 353
Db 358 IADNIR 363

```

Search completed: January 24, 2003, 19:42:43  
 Job time : 41 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 24, 2003, 19:37:39 ; Search time 14 Seconds

(without alignments)  
1054.684 Million cell updates/sec

Title: US-09-675-509-4

Perfect score: 1802  
Sequence: 1 MSTOPKTLVGLPEPLPSMN.....CVDPANKEVKNAGVLRPL 356

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	ID	Description
1	230	12.8	409 1	THI1_PANTH
2	115.5	6.4	886 1	APCE_PORPU
3	111.5	6.2	2670 1	YA05_SCHPO
4	108	6.0	885 1	APCE_AGLNE
5	104.5	5.8	512 1	AMY1_DEBOC
6	104	5.8	689 1	BCSB_PSEFL
7	101	5.6	1042 1	SVI_EORBU
8	99.5	5.5	701 1	VAB5_SCHPO
9	98.5	5.5	1160 1	C8CA_BACTP
10	97.5	5.4	1111 1	KIP1_YEAST
11	97.5	5.4	2701 1	IP3S_HUMAN
12	97	5.4	525 1	DIMH_CABEL
13	96.5	5.4	793 1	THCD_ECOLI
14	96.5	5.4	889 1	COPP_YEAST
15	96	5.3	485 1	Y045_MYCPN
16	95.5	5.3	809 1	ENPL_HORVU
17	95.5	5.3	1324 1	CUT3_SCHPO
18	95	5.3	1098 1	RPOP_MAIZE
19	94.5	5.2	394 1	LIP3_DROME
20	94.5	5.2	466 1	SVN_YERPE
21	94.5	5.2	1125 1	YE62_SCHPO
22	94.5	5.2	1868 1	YHDO_YEAST
23	94	5.2	408 1	AROA_SUITO
24	94	5.2	676 1	IF2M_YEAST
25	94	5.2	776 1	AKHE_HUMAN
26	94	5.2	1285 1	TOXA_PASMU
27	93.5	5.2	346 1	CG2A_DAUCA
28	93	5.2	346 1	PSTS_ECOLI
29	93	5.2	406 1	PKK3_HUMAN
30	93	5.2	626 1	RROC_SYNY3
31	93	5.2	1104 1	STV_YEAST
32	92.5	5.1	279 1	PHEA_LACLA
33	92.5	5.1	429 1	RNE_GUITH

34	92.5	5.1	513 1	XYLG_ECOLI	P37388 escherichia
35	92.5	5.1	662 1	TLPB_BACSU	P39217 bacillus su
36	92.5	5.1	1177 1	V307_MYCGE	P47549 mycoplasma
37	92.5	5.1	4385 1	YP73_CABEL	O09222 caenorhabdi
38	92	5.1	348 1	MO2L_ARATH	O92677 arabidopsis
39	92	5.1	551 1	YGIF_YEAST	P53214 saccharomyc
40	92	5.1	3562 1	PGCV_CHICK	O90953 gallus gall
41	91.5	5.1	376 1	MID2_YEAST	P36037 saccharomyc
42	91.5	5.1	402 1	METK_PYRAE	O82977 pyrobaculum
43	91.5	5.1	529 1	TCPB_CABEL	P47207 caenorhabdi
44	91.5	5.1	531 1	NUSA_MYCGE	P47387 mycoplasma
45	91.5	5.1	896 1	APCE_SYNY4	O02907 synechocyst

## ALIGNMENTS

RESULT 1	THI1_PANTH	STANDARD	PRT	409 AA.
AC	P45741			
DT	01-NOV-1995 (Ref. 32, Created)			
DT	01-NOV-1995 (Ref. 32, Last sequence update)			
DT	16-OCT-2001 (Ref. 40, Last annotation update)			
DE	Thiaminase I precursor (EC 2.5.1.2) (Thiamine pyridinylase).			
OC	Paenibacillus thiaminolyticus (Bacillus thiaminolyticus).			
OC	Bacteria: Firmicutes; Bacillales; Paenibacillaceae; Paenibacillus.			
OX	NCBI_TaxID=49283;			
RN	[1]			
RX	SEQUENCE FROM N.A., AND SEQUENCE OF 31-50.			
RA	Medline=96216437; Pubmed=8631946;			
RT	Costello C.A., Kelleher N.L., Abe M., McLafferty F.W., Begley T.P.;			
RT	"Mechanistic studies on thiaminase I. Overexpression and			
RT	identification of the active site nucleophile.";			
RL	J. Biol. Chem. 271:3445-3452(1996).			
RN	[2]			
RP	X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).			
RX	Medline=99060077; Pubmed=9843405;			
RA	Compagno N., Costello C.A., Begley T.P., Ealick S.B.;			
RT	"Crystal structure of thiaminase-I from Bacillus thiaminolyticus at			
RT	2.0-A resolution.";			
RL	Biochemistry 37:15981-15989(1998).			
CC	-1- FUNCTION: DEGRADATES THIAMINE BY REPLACING ITS THIAZOLE MOIETY WITH			
CC	A WIDE RANGE OF NUCLEOPHILES.			
CC	-1- CATALYTIC ACTIVITY: Thiamine + pyridine = heteropyrithiamine + 4-			
CC	methy1-5-(2-hydroxyethyl)-thiazole.			
CC	-1- SUBUNIT: MONOMER.			
CC	-1- SUBCELLULAR LOCATION: Secreted.			
CC	-1- INDUCTION: INHIBITED BY ORGANOMERCURIALS AND IODOACETATE.			
CC	-1- MASS SPECTROMETRY: MW=42127; MW_ERR=1; METHOD=Electrospray;			
CC	RANGE=31-409.			
CC	-1- MASS SPECTROMETRY: MW=42198; MW_ERR=1; METHOD=Electrospray;			
CC	RANGE=30-409.			
CC	-1- MASS SPECTROMETRY: MW=42255; MW_ERR=1; METHOD=Electrospray;			
CC	RANGE=29-409.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb.ch/announce/">http://www.isb.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; U17168; AAC44156.1; -.			
DR	PDB; 2THI; 07-OCT-98.			
DR	PDB; 3THI; 14-OCT-98.			
DR	PDB; 4THI; 14-OCT-98.			
DR	InterPro: IPR000567; SBP_bac.1.			
DR	Pfam: PF01547; SBP_bacterial_1; 1.			
KW	Transferrase; Signal; Thiamine catabolism; 3D-structure.			
FT	SIGNAL 1 29			
FT	CHAIN 30 409			
FT	THIAMINASE I.			

	FT	ACT_SITE	143	143	NUCLEOPHILE.
	FT	ACT_SITE	271	271	BASE.
	FT	VARIANT	30	30	MISSING (IN PART OF THE CHAINS).
	FT	VARIANT	31	31	MISSING (IN PART OF THE CHAINS).
	FT	MUTAGEN	143	143	C->S; INACTIVATES ENZYME.
	FT	MUTAGEN	271	271	E->Q; LOSS OF ACTIVITY.
	SQ	SEQUENCE	409 AA;	45213 MM;	9A5BD986C5C9182A CRC64;
	Query March		12.8%;	Score 230;	DB 1:
	Best Local Similarity		25.2%;	Pred. No. 5.le-09;	Length 409;
	Matches 90;	Conservative	53;	Mismatches 146;	Indels 68;
					Gaps 14;
Qy	7	TLTVGLFEPYLPSPMNGNEVKLINLIKVDLPTQVSGYNIEYTEFCDCYSADSLQSLPDVFS	66		
Dd	40	TLKVAIYPVP-----DPAHFQAALVDQMGEQGEGVKLEPFDMDSYSDAPEDDL-DVFV	92		
Qy	67	TDSIFLPLVLISG-----GVKSLSDSLVRAGVGD.LHSFYSSASAVNGSYYGFPOYLCSNF	121		
Dd	93	LDSITLSHFVDAGYLLPRPSODIDA-----EDVALPALOGAKKNGEYGGIPQLLTNL	146		
Qy	122	LSSPNGTO-QASSILEIAKVG--YEQIYVPD-----VASSSSFYFGLYOOLLOSS	171		
Dd	147	LFYRGDKLKIGVDNIYELLYKKIGTSHSQIRPPQNKGILLMAGGTTKASYMLE----	201		
Qy	172	SSANDIKAS---DLPOSQOVNKDIOTKYTIIDST-----VVASQREYINSYKOG	220		
Dd	202	--ALIDVTGQYTEYLLRPLRLPLINKVLRIGLLILNMAGEKPSGYVPBEGDDAYVASWFA	259		
Qy	221	KPISNVYVYGSESMCEIKDIRDOQVNVOLISTS---DKPYUYTPVLVALNSMLCEBKOKY	277		
Dd	260	QGSGRAFPFGYSSSWMRMGYA--EQVRFRISSAGODIPLEYSDVSVNSNSTAHNP----	313		
Qy	278	AVEVIKNTLTNYLDLGGL--TLPANKNGTAHLAKSNSFYAQISO	323		
Dd	314	--ELAKKLVANVASADTYEQLRPQADGQRYPLLPHRHQVYEALMODPVYISELAQ	368		

```

01 RESULT 2
02 APCE_PORPU STANDARD: PRT; 886 AA.
03 ID APCE_PORPU
04 AC P51263;
05 DT 01-OCT-1996 (Rel. 34, Created)
06 DT 01-OCT-1996 (Rel. 34, Last sequence update)
07 DT 01-OCT-1996 (Rel. 34, Last annotation update)
08 DE Phycobilisome linker polypeptide (Anchor polypeptide) (PBS-anchor
09 protein).
10 GN APCE.
11 OS Porphyra purpurea.
12 OC Chloroplast.
13 OX Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
14 NCBI_TaxID=2787;
15
16 (1)
17 SEQUENCE FROM N.A.
18 STRAIN=Avonport;
19 RA Reith M.E., Munholland J.;
20 RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast
21 genome." ;
22 RL Plant Mol. Biol. Rep. 13:333-335(1995).
23 CC -!- FUNCTION: THIS PROTEIN IS POSTULATED TO ACT BOTH AS TERMINAL
24 ENERGY ACCEPTOR AND AS A LINKER POLYPEPTIDE THAT STABILIZES
25 THE PHYCOBILISOME ARCHITECTURE.
26 CC -!- SIMILARITY: OF THE REPEATED DOMAINS TO N-TERMINAL REGIONS OF
27 PHYCOCYANIN ROD LINKER POLYPEPTIDES.
28
29 CC -----
30 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
31 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
32 CC the European Bioinformatics Institute. There are no restrictions on its
33 CC use by non-profit institutions as long as its content is in no way
34 CC modified and this statement is not removed. Usage by and for commercial
35 CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
36 CC or send an email to license@isb-sdb.ch).
37 CC
38 CC EMBL; U38804; AAC08149.1; -.

```

DR HSSP: P00315; 1B33.  
DR InterPro: IPR001297; PBS\_linker\_poly.  
DR InterPro: IPR01659; Phycobillismome.  
DR Pfam: PF00427; PBS\_linker\_poly; 3.  
DR Pfam: PF00502; Phycobillismome; 1.  
DR ProDom: PD000340; Phycobillismome; 2.  
KW Phycobillismome; Electron transport; Repeat;  
KW Chloroplast.  
SQ SEQUENCE 886 AA; 100160 MW; B21074C6D58C6203 CRC64;  
Query March 6.4%; Score 115.5; DB 1; Length 886;  
Best Local Similarity 20.5%; Pred. No. 1.6;  
Matches 75; Conservative 57; Mismatches 124; Indels 109; Gaps 17;

```

QY 54 SDA:LOS:LPDVFSTDS:FLPLVSLG--GVMSLDESLRGVGTDL-----HSFVS--SS 103
Db 521 SSANFENSVDIAKYV---YLRFGEVYQEEKLRLPESOLQDSQISVREFIROLAK 576
QY 104 ASVNGSVYGFPOYLCs-----NFLSSPN-GTQOASSLLELAKQVGEQIVYPDVAASS 156
Db 577 SSI:RSLTWELVLYCKAL:EVYHNRLRLRPYRGQEIINKYFDIAVKEGYQVVDALIDSPE 636
QY 157 SFTVFG-----LYOQLOLSSSAVDT--KASDLPQSCD 188
Db 637 YIEFFGENVYPERYTPPAGIALRSLRPGIIDRFKKVSSKSRFVELKVKKEIRSSND 696
QY 189 -----QYNKOITQKRYTILDS:TVVASORE---INSVQKGRISN 225
Db 697 IQSRIAGCVTALRQDSVIFDYNMSSOE---VEQALRAIYRQIFERDLSFSGIEFLD 753
QY 226 YVYVYSESMCEIKDIIND-----QOY-NNQLIGTSDK-----PYUYTDVLA 266
Db 754 IESSFLKQINVAKELIKLALSELVGEKEFYQRPYNTKVIETGKTHILGRAPNNQAEIRFL 813
QY 267 NSNL:CDEKQKAAVEVIQOLL-----TNFLVLDLGLGLTIPANKKGINHLAKSSNFYA 319
Db 814 NQIILASKGLSTFVETLVNSESVDYSGVGTNVPYRRFP--TLPA-----ANFNTETLYN 865
QY 320 QLSOO 324
Db 866 RLTKQ 870

```

RESULT 3	ID	YA05_SCHPO	STANDARD;	PRT;	2670 AA.
YA05_SCHPO	010105;				
AC	01-FEB-1996	(Rel. 33, Created)			
DT	01-FEB-1996	(Rel. 33, Last sequence update)			
DT	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	Putative translational activator C18G6.05C (GCN1 homolog).				
GN	SPAC18G6.05C.				
OS	Schizosaccharomyces pombe (fission yeast).				
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;				
OC	Schizosaccharomycetales; Schizosaccharomycetaceae;				
OC	Schizosaccharomycetes.				
OX	NCBI_TaxID=4896;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=972;				
RX	MEDLINE=21848401; PubMed=11859360;				
RA	Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,				
RA	Sgouros V., Peat N., Hayles J., Baker S., Basham D., Bowman S.,				
RA	Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,				
RA	Collins M., Connor A., Cronin A., Davis P., Felwell T., Frazer A.,				
RA	Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,				
RA	Hotiroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,				
RA	James K., Jones L., Jones M., Leather S., McDonald S., Mclean J.,				
RA	Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,				
RA	Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,				
RA	Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,				
RA	Skellon J., Simmonds M., Tivey A., Walsh S.V., Warren T., Whitehead S.				
RA	Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.				

RA Woodward J., Volckaert G., Aert R., Robben J., Grymoprez B.,  
 RA Willems I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Mosel D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Medler H., Wandt R., Pinnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rocher M., Galliard C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Arnezong J., Forburg S.L.,  
 RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shipkova G.V., Usery D., Barrell B.G., Nurse P.,  
 RT "The genome sequence of *Schizosaccharomyces pombe*."  
 RL Nature 415:871-880(2002).  
 CC -1- SIMILARITY: STRONG, TO YEAST GCN1.  
 CC -1- SIMILARITY: CONTAINS 19 HEAT REPEATS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z68198; CAA92385.1; -  
 DR InterPro; IPR000357; HEAT\_repeat.  
 DR Pfam; PF02985; HEAT\_4.  
 DR PROSITE; PS50077; HEAT\_REPEAT; 4.  
 KW Hypothetical protein; Translation regulation; Activator; Repeat.  
 FT REPEAT 315 352 HEAT 1.  
 FT REPEAT 1062 1099 HEAT 2.  
 FT REPEAT 1319 1356 HEAT 3.  
 FT REPEAT 1439 1476 HEAT 4.  
 FT REPEAT 1478 1514 HEAT 5.  
 FT REPEAT 1518 1555 HEAT 6.  
 FT REPEAT 1557 1593 HEAT 7.  
 FT REPEAT 1637 1674 HEAT 8.  
 FT REPEAT 1676 1713 HEAT 9.  
 FT REPEAT 1714 1751 HEAT 10.  
 FT REPEAT 1753 1792 HEAT 11.  
 FT REPEAT 1793 1830 HEAT 12.  
 FT REPEAT 1898 1939 HEAT 13.  
 FT REPEAT 1941 1977 HEAT 14.  
 FT REPEAT 1982 2019 HEAT 15.  
 FT REPEAT 2020 2055 HEAT 16.  
 FT REPEAT 2057 2090 HEAT 17.  
 FT REPEAT 2319 2361 HEAT 18.  
 FT REPEAT 2379 2416 HEAT 19.  
 S1 SEQUENCE 2670 AA; 297333 MW; 948E9316D56D74C3 CRC64;  
 Query Match 6.2%; Score 111.5; DB 1; Length 2670;  
 Best Local Similarity 19.7%; Pred. No. 15;  
 Matches 77; Conservative 60; Mismatches 146; Indels 107; Gaps 17;  
 QY 10 VGLF--PILP-----SNENENEVLLNL-----IDVLPVQVSGVNI 45  
 Db 1433 LGITFEPLPLPLTLTSPGDNANEVREATMDAVKQIMSQLAFGVLPLTLIDGLN- 1491  
 QY 46 EYTFDCYSDASLQSLPDVFTD--SIFLPYLVISGVKSLDESIVRGVTDGLHSFVSS 103  
 Db 1492 EYNNRSKASAEIILGMSYMAPKQSLVFLPTI-----IPKLSVLT-----DSMSQVANT 1541  
 QY 104 ASVN---GSYGRP--OYLCSNPLSSPNCGTQOASLLLELAQKVGEEQIVYPDV- 152  
 Db 1542 ANKSLRFRGDVINSPEIQTLPTLLKALSDCTRYTDALFALLKTSFVHYLDPSLALVI 1601  
 QY 153 -----ASSSFTVGLYQQLQSSSSAAV-----DIRASD 182  
 Db 1602 PILKYGRLERNAGTKRQAKIFGLMASTPEENALVYIESLMPRLREVLDPVDTAT 1661  
 QY 183 LPSGDQVNDKITQKRYTILDS--TVVASQREYINSVKGGRISNYVYGSESCE--IK 238

Db 1662 AKALGSIEIKGEKPEPTLIPELFNVRISCESEVDROGAAGSEIIAGLARLEDP 1721  
 QY 239 DIIRDQO-----YVNLIGTSDKPYVYTDVIALNSLDCDEK---QKVAVE 280  
 Db 1722 EILKNTSSPPPHIRESTISLILYLPATFGSRFPQYLARALPILSLGLADSELVQTASLR 1781  
 QY 281 VIKNLTN--TLVLDLIGLGLTTPANKGI 308  
 Db 1782 AAKMIVNNAVTKSVYDL- -LPELEKGL 1806  
 RESULT 4  
 APCE\_AGLINE  
 ID APCE\_AGLINE STANDARD; PRT; 885 AA.  
 AC P2856;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE Phycobillosome linker polypeptide (Anchor polypeptide) (PBS-anchor protein).  
 GN APCE.  
 OS Aglaetochamion neglectum.  
 OG Chloroplast.  
 OC Eukaryota; Rhodophyta; Florideophyceae; Ceramiales; Ceramiales;  
 OC Aglaetochamion.  
 OX NCBI\_Taxid=2765;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93144689; PubMed=7678762;  
 RA Apt K.E., Grosman A.R.;  
 RT "Characterization and transcript analysis of the major  
 RT phycobilliprotein subunit genes from *Aglaetochamion neglectum*  
 RT (Rhodophyta)."  
 RL Plant Mol. Biol. 21:27-38(1993).  
 CC -1- FUNCTION: THIS PROTEIN IS POSTULATED TO ACT BOTH AS TERMINAL  
 CC ENERGY ACCEPTOR AND AS A LINKER POLYPEPTIDE THAT STABILIZES  
 CC THE PHYCOBILISOME ARCHITECTURE.  
 CC -1- SIMILARITY: OF THE REPEATED DOMAINS TO N-TERMINAL REGIONS OF  
 CC PHYCOCYANIN ROD LINKER POLYPEPTIDES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z11905; CAA77957.1; -  
 DR PIR; S20621; S20621.  
 DR HSSP; P00315; 1B33.  
 DR InterPro; IPR001297; PBS\_linker\_poly.  
 DR InterPro; IPR001659; Phycobillosome.  
 DR Pfam; PF00427; PBS\_linker\_poly; 3.  
 DR Pfam; PF00502; Phycobillosome; 1.  
 DR Pfam; PD000340; Phycobillosome; 2.  
 KW Phycobillosome; Electron transport; Photosynthesis; Repeat;  
 KW Chloroplast.  
 S1 SEQUENCE 885 AA; 101153 MW; C91D75CD6307B4B9 CRC64;  
 Query Match 6.0%; Score 108; DB 1; Length 885;  
 Best Local Similarity 21.2%; Pred. No. 5.5; 119; Indels 104; Gaps 21;  
 Matches 79; Conservative 70; Mismatches 146; Indels 107; Gaps 17;  
 QY 53 YSDASLQSLPDVPS--TDSIFLP-----YLVSLGVKSLDESIVRGVTDGLHSFVSS--S 103  
 Db 518 FSDEVMSLPDDIEIGIKATILRVFGRTYVLEELSSVAKF--ESLFRSSEKISVPRIRLVK 576  
 QY 104 ASVNGSYVGGPPQYLCNSF-----LSSPN-GTQOASLLLELAQKVGEEQIVYPDVASS 156  
 Db 577 SSVFRSLYMEPLVYCKAIEYHYVRLGRPSYGRQEIINQYDIYVREGYIMIDYLLNSSE 636  
 QY 157 SFTVFG-----LYOQLQSSSSAAVDIRASDLP--PQSGDV 190



DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Cyclic di-GMP binding protein (cellulose synthase regulatory subunit).  
OS BCSB OR WSSC.  
OC Pseudomonas fluorescens.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=294;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SEM25;  
RX MEDLINE=22013850; PubMed=12019221;  
RA Spillers A.J., Kahn S.G., Bohannon J., Traviano M., Rainey P.B.;  
RT "adaptive divergence in experimental populations of Pseudomonas  
fluorescens. I. Genetic and phenotypic bases of wrinkly spreader  
phenotype."  
RT Genes 161:33-46(2002).  
RL  
CC -1- FUNCTION: Binds the cellulose synthase activator, bis-(3'-5')  
cyclic diguanylic acid (c-di-GMP) (By similarity).  
CC -1- SUBUNIT: Tightly associated with the cellulose synthase catalytic  
subunit (By similarity).  
CC -1- PATHWAY: Bacterial cellulose biosynthesis.  
CC -1- SIMILARITY: BELONGS TO THE ACSB/BCSB FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AY074776; AAJ71843.1; -  
KW Cellulose biosynthesis.  
SQ SEQUENCE 689 AA; 75272 MW; FBE0BCE76EF3D8 CRC64;  
Query Match 5.8%; Score 104; DB 1; Length 689;  
Best Local Similarity 21.4%; Pred. No. 7.4;  
Matches 60; Conservative 44; Mismatches 95; Indels 84; Gaps 12;  
QY 14 PY-LPSNNGENVEKILNLIKDVLPYOVSGYN-----I 45  
DB 267 PYDAPNMLPSNRPRVRLGELIQ-QKLSVSGYNPGASIVDMRLPDLFWNREGEVPLKLY 325  
QY 46 EYTFDCYSASLQ-SLPDVSTDSIFLPYVLSGVKSL-----DSLVRGVYTGDLHS 98  
DB 326 RYTPQOVSTNSSLIGLNDQF-MKSVALLPSVSNLGGGQGLDOLKDESLPREVT-TLLP 383  
QY 99 FVSSASVNGSVYGFPOYL-----CSNFLSSPNGTQOASSLELAOKVYEQ----- 146  
DB 384 ISSASPSKLOVRPMYDIKEGECRDIIIVDMKSSVDPDSITLV---TGQHYIAMPNIG 440  
QY 147 -----IYPPVASSSS-----FTVFGYLQQLQSSSSAAVDIKASDL 183  
DB 441 VFNDSGFPTRLADSSAVVMPDNYGTDELTAIVLIGRGCEANGPATAVKVVQAKDV 500  
QY 184 PQSGDQ-----VNKDITOKYRTILDSTVVASQREYINS 216  
DB 501 QSVADKDLVLATAANOPLKQMOQOYLPATSDGEQHOFLLS 541  
RESULT 7  
SYL BORBU  
AC SYL BORBU STANDARD; PRT; 1042 AA.  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Isolecyl-cRNA synthetase (EC 6.1.1.5) (Isoleucine--cRNA ligase)  
(1Iers).  
DE ILES OR BB0833.  
OS Borrelia burgdorferi (Lyme disease spirochete).

OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 35210 / B31;  
RX MEDLINE=98065943; PubMed=9403685;  
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,  
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,  
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,  
RA Peterson J., Kertavage A.R., Quackenbush J., Salzberg S., Hanson M.,  
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,  
RA Uterback T., Wathley L., McDonald L., Atlach P., Bowman C.,  
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,  
RA Smith H.O., Venter J.C.;  
RT "Genomic sequence of a Lyme disease spirochete, Borrelia  
burgdorferi."  
RT Nature 390:580-586(1997).  
RL  
CC -1- CATALYTIC ACTIVITY: ATP + L-isoleucine + tRNA(ile) = AMP +  
diphosphate + L-isoleucyl-tRNA(ile).  
CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).  
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AE001181; AAC67179.1; -  
DR HSPB; P56690; 11LE.  
DR TIGR; BB0833; -  
DR InterPro: IPR002300; tRNA-synt\_1a.  
DR InterPro: IPR001412; tRNA-synt\_1.  
DR InterPro: IPR002301; tRNA-synt\_1le.  
DR Pfam: PF00133; tRNA-synt\_1; 1.  
DR PRINTS; PRO0984; TRNASYNTHILE.  
DR TRFAM; TIGR00392; 11leS; 1.  
DR POSITIVE; PS00178; AA\_TRNA\_LIGASE\_1; FALSE\_NEG.  
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
KW Metal-binding; zinc; Complete proteome.  
FT SITE 48 58 "HIGH" REGION.  
FT SITE 594 598 "KMSK" REGION.  
FT BINDING 597 597 ATP (BY SIMILARITY).  
SQ SEQUENCE 1042 AA; 122331 MW; 6C0F7D820CA32F75 CRC64;  
Query Match 5.6%; Score 101; DB 1; Length 1042;  
Best Local Similarity 18.8%; Pred. No. 22;  
Matches 74; Conservative 59; Mismatches 120; Indels 140; Gaps 16;  
QY 19 WNEGNENVEKILNLIKDVLPYOVSGYNIEYTFDCYSASLQSLPDVSTDSIFLPYVLSL 78  
DB 634 YSDNG-----VRDLK-----NIIIPINWASFTYIALLDKRP-----PKMLSL 674  
QY 79 GGVKSLDESIVRGVTDGLHSFVSSASVNGSVYGFPOYLCNFLSSPNGTQOASSLEL 138  
DB 675 AKNNMLDKWII-----SELSELKK-----ILNTEIDKYNLTGSIESTLEF 714  
QY 139 AKYV-----GYEQIYVPD-----VASSSFTVFGIYQDLQ 169  
DB 715 IDKLNWYIRSRRRRPFKSNQDKNDAYETLLYALITMLILAPLPITTEIYQNLKT 774  
QY 170 SSSSAVDIKASDLPOSQDQ-VNKDITOKYRTILDSTVVASQREYINSVQKQPISNVYV 228  
DB 775 DEBKQSHL--NDYPRANENFINKTIEKINLARKITSMARSLRLNINIRNPISITII 832  
QY 229 -----GVSBMCEIKDIIRDQ-----QYNVOLIGT--SDKPYVY 260  
DB 833 VTKQNQNNMLMEMQETIILDEINAKEMKIKANEELITYYAKANFELGKGLKDKMAVS 892

OY 261 TDV-----LATNSLDEKOKAVEYIK-----N 284  
DB 893 TEISLKNEDIIKIINGSYEIKANAHYLSLNDIILEREKENLKVINEESTITIGID 952  
OY 285 LITNTVLDTGLGLTPANKNGIAHLAKSNF 317  
DB 953 LITKELIYE-----GLTREFVRO-IONLAKENKF 980  
RESULT 8  
YABS\_SCHPO STANDARD: PRT: 701 AA.  
AC 009807;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DE 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Hypothetical protein C2G11.05c in chromosome I.  
GN SPAC2G11.05c.  
OS Schizosaccharomyces pombe (fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_Taxid=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Gwilliam M., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Aguires J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA Jones K., Jones L., Jones M., Leach S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skellern J., Simmonds M., Squares R., Squares S., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Voiclaert G., Aert R., Robben J., Grymoult B.,  
RA Wellens J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer T., Beck A., Leirich H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Huret S.M.,  
RA Lucas M., Rocher M., Galliard C., Tallada V.A., Garzon A., Rhode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Snpakovski G.V., Ussery D., Bartell B.G., Nurse P.;  
RT "The genome sequence of Schizosaccharomyces pombe.";  
RL Nature 415:871-880(2002).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: Z54354; CA91170.1; .  
DR InterPro: IPR004328; BRO1.  
DR Pfam: PF03097; BRO1; 1.  
KM Hypothetical protein.  
SQ SEQUENCE 701 AA; 81769 MW; E13BE4CAAE085671 CRC64;  
Query Match 5.5%; Score 99.5; DB 1; Length 701;  
Best Local Similarity 19.0%; Pred. No. 16;  
Matches 60; Conservative 68; Mismatches 133; Indels 55; Gaps 13;  
OY 4 OPTILTVGLFPYLPBSEWNGEVLKILIKOVLPQVSGNIEYTEPDCYSASLQSLPD 63

DB 309 QPR-IPVGSFPLDT-----VTMVGSIPTDVLDTXATAPYSTCLDEPFMDLE 357  
OY 64 VFTSDSIFLPYLVISGVKSLDESLVGVGTDLHSFVSSASVNGSVY--GFPQYLCSNF 121  
DB 358 EHEKN-----LIDAAVKSID-CLTKEGTBKLPALMEELKSVNDKNYIETSKWMENF 409  
OY 122 L-ISSPNGTQ---QASLLLEAKV-----GYEQIVPDPVASSSFTYFGLYQOL 167  
DB 410 TEIKNLGLGLEFRSEASSLCLIDKNGVTRHCCQALDEII-NSVKSQNMKEKGFYSDV 468  
OY 168 LQSSSSAA-VDIKASDLPQSGDQVNDITQKRTILDTSTVASQREYINSVKQKPISNY 226  
DB 469 IKLHEEVSLLKVENVSQNAKDITIQKDISAGKDFELS-AQSERILSNLAEKGNPTDQ 527  
OY 227 YVGSSEMECKIDIRDOQVNVQITGSDPYV-----YTDVLA-----NSN 269  
DB 528 LILQANVYLQWDLKDERNMKTKLSGNDYFNISMYSNHTDHSLLKCFREAVDAKKNFR 587  
OY 270 LCKEKQKVAEVIKNL 285  
DB 588 RQREORIEIIONVEKL 603  
RESULT 9  
C8CA\_BACTP STANDARD: PRT: 1160 AA.  
ID C8CA\_BACTP  
AC 045706;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Pesticidal crystal protein cry8Ca (insecticidal delta-endotoxin  
DE CryVIIIcA) (crystalline entomocidal protoxin) (130 kDa crystal  
DE protein).  
GN CRY8CA OR CRYVIII(C) OR CRYVIII.  
OS Bacillus thuringiensis (subsp. japonensis).  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_Taxid=128936;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=bulbul;  
RX MEDLINE=94100786; PubMed=7764305;  
RA Sato R., Takeuchi K., Ogiwara K., Minami M., Kaji Y., Suzuki N.,  
RA Hori H., Asano S., Ohba M., Iwahana H.;  
RT "Cloning, heterologous expression, and localization of a novel crystal  
RT protein gene from Bacillus thuringiensis serovar japonensis strain  
RT bulbul toxic to scarabaeid insects.";  
RL Curr. Microbiol. 28:15-19(1994).  
RN [2]  
RP SEQUENCE OF 1-14 AND 56-64, AND CHARACTERIZATION.  
RX MEDLINE=94259659; PubMed=8200856;  
RA Hori H., Suzuki N., Ogiwara K., Himejima M., Indrasith L.S.,  
RA Minami M., Asano S., Sato R., Ohba M., Iwahana H.;  
RT "Characterization of larvicidal toxin protein from Bacillus  
RT thuringiensis serovar japonensis strain bulbul specific for  
RT scarabaeid beetles.";  
RL J. Appl. Bacteriol. 76:307-313(1994).  
CC -1- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT  
CC EPITHELIAL CELLS OF INSECTS. ACTIVE ON VARIOUS SCARABAEID BEETLES  
CC SUCH AS ANOMALA CUPREA, A. RUFOCUPREA AND POPILLIA JAPONICA.  
CC -1- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING  
CC SPORELIATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART OF  
CC THE SPORE COAT.  
CC -1- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE N-  
CC TERMINUS.  
CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

```

CC or send an email to license@sb-sib.ch).
CC -----
DR EMBL: U04366; AAA2119.1; -.
DR HSSP: P07130; 1DLC.
DR InterPro: IPR001178; Endocoxin.
DR Pfam: PF00555; endotoxin; 1.
DR Tricin; Sporulation.
KM SEQUENCE 1160 AA; 130425 MW; C16C3D912EBB8751 CRC64;
SQ
Query Match 5.4%; Score 98.5; DB 1; Length 1160;
Best Local Similarity 22.6%; Pred. No. 38;
Matches 91; Conservative 58; Mismatches 126; Indels 127; Gaps 25;
QY 26 VKLNLKDVLPVSGN-----IEYTFD-CYSDASIQSPDYSTSIPLP-YLV 76
DQ 728 VQANLLQDTGFNRINNGENGTGTGIEVEGVDFLFXKRSIR-LTSAEIDETETPTL- 785
QY 77 SLGKSLDESLVGVG-DLHGFVSSASV-----NGSYGFP----- 114
DQ 786 ---YQIDESLKPYYRYKLGFGISQDLEIKLIRHRAQIVKVPDNLPPVRPVNS 841
QY 115 -----QYLCNFFLL--SSPNGTQASS-----LLELAQKVGEOIVVYPDV 152
DQ 842 CGGVDRSGEOQYVDANLALENNGENKNSSDSHAFPHITGEIDLENNTGI-MIVFKIP 900
QY 153 ASSSFTVFGLYQQLLOSSSAVDIKASDLPQSGDQVN-KDITQKRYTLDSTVVASQ 210
DQ 901 TTNGNATLGNL-EFVEEG-----PLSGETLEMAQQEQEQMDKMRKRAASE 946
QY 211 REYINSYKQKGPISNYVYGE---SKCEIKDIIIDQOVVOLIGTSKPYVTVDV- 264
DQ 947 KTYV-AAKQA--IDRLFADYQDKLNSEVMSDILAAQNL-VQSI-----PYVNDALPE 997
QY 265 --AANSNLCEKQKVAEVIKNLTLNL-VLDLGLGLTIPAN-KNGIHLAKSS--- 315
DQ 998 IPGNNVYSFTE-----LTNRQQQWNLVDLQNALPNCDFRGLSNMNTSDVN 1045
QY 316 -----NRYAQLSQCFDAKESSEVLRCDVPFANKE 344
DQ 1046 VQQLSDTSVLVTPMNSGVSGQFTVQPPRYRVLVT--ARKE 1085
RESULT 10
KIP1_YEAST STANDARD; PRT; 1111 AA.
ID KIP1_YEAST
AC P28742;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
D 15-JUL-1998 (Rel. 36, Last annotation update)
D Kinesin-like protein KIP1.
GN KIP1 OR CIN9 OR YBL063W OR YBL0504 OR YBL0521.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=49132;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=94205266; PubMed=8154187;
RA Scherens B., el Bakoury M., Vierendeels F., Dubois E., Messenguy F.;
RT "Sequencing and functional analysis of a 32,560 bp segment on the
RT left arm of yeast chromosome II. Identification of 26 open reading
RT frames, including the KIP1 and SECI7 genes.";
RN Yeast 9:1355-1371 (1993).
RN [3]

```

```

RP CHARACTERIZATION.
RC STRAIN=S288C;
RX MEDLINE=92354062; PubMed=1643659;
RA Saunders W.S., Hoyt M.A.;
RT "Kinesin-related proteins required for structural integrity of the
RT mitotic spindle.";
RL Cell 70:451-458 (1992).
CC -1- FUNCTION: REQUIRED FOR ASSEMBLY OF THE MITOTIC SPINDLE. INTERACT
CC WITH SPINDLE MICROTUBULES TO PRODUCE AN OUTWARDLY DIRECTED
CC FORCE ACTING UPON THE POLES. FOLLOWING SPINDLE ASSEMBLY, CIN8 AND
CC KIP1 APPARENTLY ACT TO OPPOSE A FORCE THAT DRAWS SEPARATED POLES
CC BACK TOGETHER. THIS FORCE SEEMS TO BE MEDIATED BY KAR3.
CC -1- SUBUNIT: MIGHT BE DIMERIC.
CC -1- SUBCELLULAR LOCATION: SPINDLE MICROTUBULES THAT LIE BETWEEN THE
CC POLES.
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. B1MC
CC SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.sib-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: Z11962; CAA78019.1; -.
DR EMBL: Z33261; CAA80785.1; -.
DR EMBL: Z35824; CAA84883.1; -.
DR PIR: A42640; A42640.
DR HSSP: P17119; 3KAR.
DR SGD: S0000159; KIP1.
DR InterPro: IPR002570; Hpt.
DR InterPro: IPR001752; kinesin_motor.
DR Pfam: PF00225; kinesin_1.
DR PRINTS: PR00380; KINESINHEAVY.
DR SMART: SM00073; HPT; 1.
DR SMART: SM00129; KISC; 1.
DR PROSITE: PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE: PS00411; KINESIN MOTOR DOMAIN2; 1.
DR PROSITE: PS00657; KINESIN MOTOR DOMAIN2; 1.
KW Motor protein; Cell division; Microtubules; ATP-binding; Coiled coil;
KW Mitosis; Cell cycle.
FT DOMAIN 50 423 KINESIN-MOTOR (BY SIMILARITY).
FT DOMAIN 424 510 COILED COIL (POTENTIAL).
FT DOMAIN 648 670 COILED COIL (POTENTIAL).
FT DOMAIN 710 780 COILED COIL (POTENTIAL).
FT DOMAIN 808 828 COILED COIL (POTENTIAL).
FT NP BIND 141 148 ATP (BY SIMILARITY).
SQ SEQUENCE 1111 AA; 125794 MW; 212F8279766137FC CRC64;
Query Match 5.4%; Score 97.5; DB 1; Length 1111;
Best Local Similarity 21.8%; Pred. No. 42;
Matches 71; Conservative 50; Mismatches 112; Indels 93; Gaps 15;
QY 24 NEVKLNLK---KDVLPVSGVNIETFEPCYSD-----ASIQSLPV-----PSTDS 69
DQ 489 NQDDINLLQSEKELAIQINPNVDFSNF--YSEIQKHTLLELMEVETIQORDSLEN 546
QY 70 IFPPYLV-----SLGVKSLDESLVRCVTDGLHGFVSSASVNG 108
DQ 547 SQKQVNTNQWQKISQVQLQTLNTLQSGSLNNYNSKSEVYKVTETLTNNVTHAKHD 606
QY 109 SVYGFQYLCNFFLSPKNTQQAASLLELAQKVGEOIVVYPDVASSSTFVGLVQQL 168
DQ 607 STLSKLTNTTLLNMQW-----ELVRSI-----STSLEIF----- 638
QY 169 QSSSAAVDIKASDLPQSGDQVKNKITQKRYTLDSTVVASQREYINSYVQKRP----- 222
DQ 639 QSDSTSHYRDLNLEIYQSHQFLKNIQNDKSLDLS--IGSILTSINEISQNTLNSM 697
QY 223 ---ISNYVYGS---ESMCEIK---DIIRDQ---QYVQQLIGTSKPYVTVDLALN 267
DQ 698 NVLIENQSGSKLKEQDLEIKLKNLNLINERRISGNFQQL--AEMKRYFDHVSRT 755

```

QY 268 SNLCDEKQKVAEVIKMLINTLVLD 293  
Db 756 SEFHDELNK-----CIDMLKDKQSKLD 777

RESULT 11  
IP3S\_HUMAN STANDARD; PRT; 2701 AA.

AC Q14571; O94773;  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Inositol 1,4,5-trisphosphate receptor type 2 (Type 2 inositol 1,4,5-trisphosphate receptor) (Type 2 Insp3 receptor) (IP3 receptor isoform 2) (Insp3R2).  
GN ITPR2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
NCBI\_TaxID=9606;  
RA [1]  
RP SEQUENCE FROM N.A. (LONG ISOFORM).  
RX MEDLINE=9436219; PubMed=8081734.  
RA Yamamoto-Hino M., Sugiyama T., Hikici K., Mattei M.-G., Hasegawa K., Sekine S., Sakurada K., Miyawaki A., Furuchi T., Hasegawa M., Mikoshiba K.;  
RT "Cloning and characterization of human type 2 and type 3 inositol 1,4,5-trisphosphate receptors."  
RL Recept. Channels 2:9-22(1994).  
RN [2]  
RP SEQUENCE FROM N.A. (SHORT ISOFORM).  
RC TISSUE=Heart;  
RX MEDLINE=98399819; PubMed=9729462;  
RA Futatsugi A., Kuwajima G., Mikoshiba K.;  
RT "Muscle-specific mRNA isoform encodes a protein composed mainly of the N-terminal 175 residues of type 2 Ins(1,4,5)P3 receptor."  
RL Biochem. J. 334:559-563(1998).  
CC -1- FUNCTION: RECEPTOR FOR INOSITOL 1,4,5-TRISPHOSPHATE, A SECOND MESSENGER THAT MEDIATES THE RELEASE OF INTRACELLULAR CALCIUM.  
CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic reticulum.  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A SHORT FORM/ITPR; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- TISSUE SPECIFICITY: THE SHORT ISOFORM IS FOUND IN SKELETAL MUSCLE AND HEART.  
CC -1- DOMAIN: THE RECEPTOR CONTAINS A CALCIUM CHANNEL IN ITS C-TERMINAL EXTREMITY. ITS LARGE N-TERMINAL CYTOPLASMIC REGION HAS THE LIGAND-BINDING SITE IN THE N-TERMINUS AND MODULATORY SITES IN THE MIDDLE PORTION IMMEDIATELY UPSTREAM OF THE CHANNEL REGION.  
CC -1- PPM: PHOSPHORYLATED ON TYROSINE RESIDUES (BY SIMILARITY).  
CC -1- MISCELLANEOUS: CALCIUM APPEARS TO INHIBIT LIGAND BINDING TO THE RECEPTOR. MOST PROBABLY BY INTERACTING WITH A DISTINCT CALCIUM-BINDING PROTEIN WHICH THEN INHIBITS THE RECEPTOR.  
CC -1- SIMILARITY: BELONGS TO THE INSP3 RECEPTOR FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announce/](http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL, D26350; BAA0384.1; -  
DR EMBL, AB012610; BAA33961.1; -  
DR Genew; HGNC:6181; ITPR2.  
DR MIM; 600144; -  
DR InterPro; IPR000699; Ca-rel channel.  
DR InterPro; IPR001682; Ca/Na\_pore.  
DR InterPro; IPR000493; Insp3\_receptor.  
DR InterPro; IPR000636; M-channel\_nlg.

DR InterPro; IPR003608; MIR.  
DR Pfam; PF00520; Ion\_trans; 1.  
DR Pfam; PF01365; RyR\_ITPR; 2.  
DR Pfam; PF02815; MIR; 4.  
DR PRINTS; PR00779; INSP3RECEPTR.  
DR SMART; SM00472; MIR; 4.  
KW Receptor; Transmembrane; Glycoprotein; Phosphorylation; Endoplasmic reticulum; Ionic channel; Ion transport; Calcium channel;  
KW Alternative splicing.  
FT DOMAIN 1 2227 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 2228 2248 POTENTIAL.  
FT TRANSMEM 2249 2260 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 2261 2281 POTENTIAL.  
FT DOMAIN 2282 2307 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 2308 2328 POTENTIAL.  
FT DOMAIN 2329 2351 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 2352 2372 POTENTIAL.  
FT TRANSMEM 2373 2394 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 2395 2415 POTENTIAL.  
FT DOMAIN 2416 2521 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 2522 2542 POTENTIAL.  
FT TRANSMEM 2543 2701 CYTOPLASMIC (POTENTIAL).  
FT MOD RES 2607 2607 PHOSPHORYLATION (POTENTIAL).  
FT VARSPPLIC 176 181 IYVGDK -> DASFWI (IN SHORT ISOFORM).  
FT VARSPPLIC 182 2701 MISSING (IN SHORT ISOFORM).  
SQ SEQUENCE 2701 AA; 308074 MW; EB5C7DDCD1F74A CRC64;

Query Match 5.4%; Score 97.5; DB 1; Length 2701;  
Best local Similarity 22.6%; Pred. No. 1.5e+02;  
Matches 91; Conservative 68; Mismatches 131; Indels 113; Gaps 23;

QY 5 PKLTVLGFLPYLPSPWNGN---EVKLINLI-----KDLPTQVSGVNIETEF 50  
Db 1371 PLAVHITLVELLAACGKGVYTEIKNSLLPLDIDIVRVVTHDDCIPE---VKIAVNF 1426

QY 51 --DQYASASQSLPDVSTDSI---FLPYVSLGV-----KSLDESIVR 90  
Db 1427 VNHCVYPTVEE-MKEITSHNIWKLFPNPLVDMAVGVNTTDRGHADIFLEKVTESIMN 1485

QY 91 GVTGDLIS-FVSSASVNGSVGFGPYLCNFIILSS---FNGTQOAS-----SLBLAQ 140  
Db 1486 IVSFFPSPSDNSTSIQTHQPVFIQLQSAFRIYNTCTWNPQAKAVESCIRLAEVAK 1545

QY 141 KVGEGQIVPDPVAVSSSFYVGLYQQLQSSSS---AAVDIKASDLPQSGDQVKNDI-- 194  
Db 1546 NRG---IAIPVDLDSQVNTLF-----MKSHSNVWGQAAGMRIS--ARSGPRFEKALG 1594

QY 195 -TKKRTILD--STVVASQREYINSVQCKPISNYYVYGESEMC--RIKDI-IRDOQYV 248  
Db 1595 PAMDYRNIEKLDQDVASLSHQFSPMMQAEFSVLVDVLSPELLFPGSDARIRCGAFWS 1654

QY 249 QLTGTSKPPYVTVLALNSNLCDKQKVAEVIKMLLT-----NTLVLDL- 295  
Db 1655 KLINHTKK-----LMEKEKLCIKIQLITREMLEKDSVVEGNTLRKILN 1701

QY 296 -----GLGTLTPANKNGIAHLAKSSNFYAQLSQQDAKASE 331  
Db 1702 RYFKGDYSIGV-----NG--HLSGAVSKTAQVGVSGSDSD 1736

RESULT 12  
D1MH\_CABEL STANDARD; PRT; 525 AA.

ID D1MH\_CABEL  
AC 017397;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Diminuto-like protein.  
GN F52H2.6.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
OC Rhabditidae; Pelodolerinae; Caenorhabditis.  
NCBI\_TaxID=6239;

```

RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Bristol N2.
RA      Chisoe S., Hawkins J.;
RL      Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC      -1- SIMILARITY: BELONGS TO THE DIMITO FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on items
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sb.ch/announce/
CC      or send an email to license@isb-sb.ch).
CC      -----
DR      EMBL; AF026214; AAB71310.1; -
DR      Wormpep; F5282.6; GI10874.
DR      InterPro: IPR001575; Oxid_FAD_bind.
DR      Pfam; PF01565; FAD_binding_4; 1.
DR      Kyr Hypothetical protein.
RY      SEQUENCE 525 AA; 61421 MW; F4D567CCE2825928 CRC64;
S1      -----
Query Match          5.4%; Score 97; DB 1; Length 525;
Best Local Similarity 20.1%; Pred. No. 16;
Matches 61; Conservative 50; Mismatches 91; Indels 102; Gaps 14;
QY      16 LPSNNGNGENVLLINLIKDVLPYQSGVNIETEPDCYSDASLQSLPVPFSTDSIF--- 71
DB      73 LKENNDNRKSKSLVN-----ARPGMTMSFRPLYKENTK-----IATDKLPDILD 119
QY      72 -----LPLVSLG-----GVKSLDESLYRGVTGDLHSFVSSASVANG 108
DB      120 LDVEKMTVAEEGVTMGQLQYLIRSGYTLPLPELDLDTYVGL-----INGGVESG 172
QY      109 SV-YGFPQYLGNTFLSSNG-----TQASLLELAQKGYEQ----- 146
DB      173 SFKYMFGHICGVYLVVMSDGLKRVYDPDSAKTEQAKQDNSLPFALPWSQGTICFLVAA 232
QY      147 ---IV-----YPPVASSSFTVFGLYQQLDS--SSAAYDI-----KAS 181
DB      233 TIKIIPCKKYVLTLYKKTETLTSEKCEQLTDESDRNSENDPFEALMFKKCKITLGEPS 292
QY      182 DLPGSGDVNKKDITQKYRTILDSIV---VASGREYINSYKQKGPISNYYGVSESM-CEI 237
DB      293 DGPDRHDEYVNPVIGRWYKMKFYTHVEDLTKKGESIEYI---PLRDYVHRRSKSIPMEL 348
QY      238 KDII 241
DB      349 RDIV 352
S1      -----
RESULT 13
YHCD_ECOLI
ID      YHCD_ECOLI STANDARD; PRT; 793 AA.
AC      P45420;
DT      01-NOV-1995 (Rel. 32, Created)
DT      01-NOV-1995 (Rel. 32, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Hypothetical outer membrane usher protein yncD precursor.
DS      YHCD OR B3216.
OS      Escherichia coli.
OC      Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC      Escherichia.
OX      NCBI_Taxid=562;
[1]
RN      SEQUENCE FROM N.A.
RC      STRAIN=K12 / MG1655;
RX      MEDLINE=9742617; PubMed=9278503;
RA      Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA      Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA      Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA      Mau B., Siao Y.;
RT      "The complete genome sequence of Escherichia coli K-12.";
```

[illegible]

```

RC STRAIN=RSY255;
RX MEDLINE=95014199; PubMed=7929113;
RA Duden R., Hosobuchi M., Hamamoto S., Winey M., Byers B., Schekman R.;
RT "Yeast beta- and beta'-coat proteins (COP). Two coatomer subunits
RT essential for endoplasmic reticulum-to-Golgi protein traffic.";
RL J. Biol. Chem. 269:24486-24495(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=8288c / FY1679;
RX MEDLINE=96437978; PubMed=8840506;
RA Bacciaro V., Eraso P., Portillo F., Mazon M.J.;
RT "Sequence analysis of a 14.6 kb DNA fragment of Saccharomyces
RT cerevisiae chromosome VII reveals SEC27, SSM1b, a putative
RT S-adenosylmethionine-dependent enzyme and six new open reading
RT frames.";
RL Yeast 12:887-892(1996).
RN [3]
RP SEQUENCE OF 1-27.
RC STRAIN=BU926;
RX MEDLINE=94009692; PubMed=8405452;
RA Harter C., Draken E., Lottspeich F., Wieland F.T.;
RT "Yeast coatomer contains a subunit homologous to mammalian
RT beta'-COP.";
RL PBBS Lett. 332:71-73(1993).
CC -1- FUNCTION: THE COATOMER IS A CYTOSOLIC PROTEIN COMPLEX THAT BINDS
CC TO DILYSINE MOTIFS AND REVERSIBLY ASSOCIATES WITH GOLGI NON-
CC CLATHRIN-COATED VESICLES, WHICH FURTHER MEDIATE BIOSYNTHETIC
CC PROTEIN TRANSPORT FROM THE ER, VIA THE GOLGI UP TO THE TRANS GOLGI
CC NETWORK. COATOMER COMPLEX IS REQUIRED FOR BUNDLING FROM GOLGI
CC MEMBRANS, AND IS ESSENTIAL FOR THE RETROGRADE GOLGI-TO-ER
CC TRANSPORT OF DILYSINE-TAGGED PROTEINS.
CC -1- SUBUNIT: OLIGOMERIC COMPLEX THAT CONSISTS OF AT LEAST THE ALPHA,
CC BETA, BETA', GAMMA, DELTA, EPSILON AND ZETA SUBUNITS.
CC -1- SUBCELLULAR LOCATION: THE COATOMER IS CYTOPLASMIC OR POLYMERIZED
CC ON THE CYTOPLASMIC SIDE OF THE GOLGI, AS WELL AS ON THE
CC VESICLES/BUDS ORIGINATING FROM IT (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U11237; AAA61711.1; -
DR EMBL: X92670; CAA63359.1; -
DR EMBL: Z72659; CAA96848.1; -
DR SGD: S0003105; SEC27.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 6.
DR PRINTS: PR00320; GPROTEINERPT.
DR ProDom: PD000018; WD40; 3.
DR SMART: SM00320; WD40; 6.
DR PROSITE: PS00678; WD REPEATS_1; 1.
DR PROSITE: PS50082; WD REPEATS_2; 5.
DR PROSITE: PS50294; WD REPEATS_REGION; 1.
DR Transport: Protein transport: Golgi stack; Membrane; Repeat;
DR WD repeat; Endoplasmic reticulum.
FT REPEAT 11 41 WD 1.
FT REPEAT 53 83 WD 2.
FT REPEAT 95 125 WD 3.
FT REPEAT 138 169 WD 4.
FT REPEAT 182 214 WD 5.
FT REPEAT 226 256 WD 6.
SQ SEQUENCE 889 AA; 99444 MW; 6A5E50BBE02CB58 CRC64;
Query Match 5.4%; Score 96.5; DB 1; Length 889;
Best Local Similarity 19.8%; Pred. No. 36;
Matches 65; Conservative 56; Mismatches 127; Indels 79; Gaps 14;
QY 19 WNEGNENKVLINLIKDVLPQTQVSGYNIEYTEPDCYSDASIQ--SLPDVFSTDSIFLPYLVS 77

```

```

DB 476 WSDNGELWMTVNTNSN--GDEASGYTLLENK-DAYLEAANNINDDSEGVDEAF----- 526
QY 78 IGGVKSILDESIVRGV-TGDLHSFVSSASVYV-----GSVGFPOYLCSNPLSSPNGTQA 132
DB 527 -DVLYELSESITSGKMWGDVFIPTATNRLNRYFVGKTYVLAHYTKMYLL-----GYLAR 581
QY 133 SLLLELAQK---VGYEQIYPDVASSSFTVFGLYQQLQ-----SSSAVADI 178
DB 582 DNKRYLLADREVHYGYE--ISLEVLPEQTLTLRGEIEEALENVLPVEGKDSLTKLARFL 639
QY 179 KASDLPQSGQVQVKDITQKRTILDSITVVASQREYINSVQGKPISNVYVYGESMEIK 238
DB 640 EGGYEIEALNISPDQDFELAL-----KVQ-----LTLAR 672
QY 239 DIIRDQVYNQILIGTSDPKPYVTYDVLANSNLCDKQKVAVEYIKMLTNTLVLDLGLG 298
DB 673 DLITDESAEMKRALGD-----ASLGRFNFKLAVEA-----FTNAHDLBSJFL 716
QY 299 LTLPAKNGIAHLAKSSNFYAQLSQQFPA 327
DB 717 HSSFNNKEGLVTLAKDERAKFENLAFA 745
RESULT 15
Y045_MYCPN STANDARD; PRT; 485 AA.
AC P75056;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical lipid-protein MG045 homolog precursor (D09_orf485).
GN MG045 ORF MG096.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfreisch R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Hermann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000011; AAB95744.1; -
DR InterPro: IPR000044; Lipoprct MG045.
DR Pfam: PF02030; Lipoprotein_8; 1.
DR PRINTS: PR00905; MYCMG045.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
DR Hypothetical protein; Lipoprotein; Membrane; Signal;
DR Complete proteome.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 485 HYPOTHETICAL LIPOPROTEIN MG045 HOMOLOG.
FT LIPID 23 23 N-ACYL DIGLYCERIDE (POTENTIAL).
SQ SEQUENCE 485 AA; 55019 MW; 1AD4E0AB211B64F8 CRC64;
Query Match 5.3%; Score 96; DB 1; Length 485;
Best Local Similarity 21.8%; Pred. No. 16;
Matches 61; Conservative 47; Mismatches 116; Indels 56; Gaps 13;
QY 24 NEVKLNLIKDVLPQTQVSGYNIEY--TEPDCYSDASIQ--SLPDVF--STDSIFLPYLVS 77

```

```

Db      201 HEVK-NTYDVNPT---GSTLNYFGVYESFANIGLKRDNLTILFVNSDSNIIINELAN 255
QY      78 ---LGVKSJDESIVRGVTDLHSHFVSSASVNGSVYGFPOYLCSNPLSSPNGTQOASS 134
Db      256 GRROGGIVYNGDAVYAAIGDLDLDEINENNL PNG-----DNFHIYQPKHSPVALD 305
QY      135 LLELAQKVGEQIVYPVASSSFTVFGLYOQLLOSSSSAAVDIKASDLPQSGDQVVKDI 194
Db      306 FLINQ---QTHFRDAA-----HQLIYQLALBEGADQTAEBELKTDEKGTSD 350
QY      195 TOKKRTILDSTVVASOR-EYINSVKQKGPISNYVGY--SESMCEIKDITRDQYVQL 250
Db      351 EDDY-----TYGAMQNFSSYVNVYSPILKNISDEFTGIVFKENKQADTKQVVKOOSOSEQ 404
QY      251 IGTSKDP-----YVTDVLALNSNLCDEKQKVAVEYIK 283
Db      405 SESAKETEEDDFTYATLKSLLKADSLDDKAKKLVDTIK 444

```

Session completed: January 24, 2003, 19:41:59  
 Job time: 18 secs

PAGE LEFT BLANK

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 24, 2003, 19:40:35 ; Search time 22 seconds

(without alignments)  
1555.630 Million cell updates/sec

Title: US-09-675-509-4

Perfect score: 1802

Sequence: 1 MSTRQPKTLTVGLFEPYLPSPWN.....CVDPAKKEVKNAGVLRPPL 356

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73:.\*  
1: p1r1:.\*  
2: p1r2:.\*  
3: p1r3:.\*  
4: p1r4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	230	12.8	409	2 T47118	thiamine pyridinyl
2	116	6.4	1248	2 B96827	hypochlorite prote
3	115.5	6.4	886	2 S73184	phycobiosome link
4	111.5	6.2	613	2 B90294	hypochlorite prote
5	109	6.2	2670	2 T37919	GCN1 homolog - fis
6	108.5	6.0	448	2 T06799	fimbria/plastin-11
7	108	6.0	998	2 T32787	hypochlorite prote
8	108	6.0	528	2 T22583	hypochlorite prote
9	108	6.0	885	2 S20621	phycobiosome anch
10	106.5	5.9	1033	2 A96714	hypochlorite prote
11	106.5	5.9	2819	2 A90551	conserved hypochl
12	106	5.9	545	2 D75208	sugar abc transpor
13	106	5.9	1054	2 G82934	hypochlorite prote
14	105.5	5.9	358	2 AB1746	hypochlorite prote
15	105.5	5.9	1277	2 E70224	hypochlorite prote
16	104.5	5.8	503	2 S06115	benzoylformate dec
17	104.5	5.8	512	2 E90325	alpha-amylose (EC
18	104.5	5.8	628	2 A82889	hypochlorite prote
19	103.5	5.7	512	2 S23355	alpha-amylose (EC
20	103.5	5.7	682	2 G90284	hypochlorite prote
21	103	5.7	286	2 A86665	outer membrane lip
22	103	5.7	637	2 AH1519	hypochlorite prote
23	102.5	5.7	685	2 G69864	penicillin-binding
24	101.5	5.6	728	2 D86278	hypochlorite prote
25	101.5	5.6	831	2 F90294	hypochlorite prote
26	101.5	5.6	1787	2 AG1360	probable tape-meas
27	101.5	5.6	1787	2 D97316	probable S-layer p
28	101.5	5.6	2401	2 T28676	thoptry protein -

30	101	5.6	1042	2 H70203	isooleucine-tRNA 11
31	100.5	5.6	745	2 E64559	outer membrane pro
32	100.5	5.6	937	2 B86210	protein F22G5.6 [1
33	100	5.5	433	2 S77508	sera protein - Syn
34	100	5.5	672	2 D81746	type III secretion
35	99.5	5.5	701	2 S62460	hypochlorite prote
36	99	5.5	507	2 S33921	alpha-amylose (EC
37	99	5.5	566	2 F70028	transmembrane rece
38	99	5.5	889	2 T45691	receptor-like prot
39	99	5.5	1159	2 S22768	130K protein - mal
40	99	5.5	1657	2 T25421	hypochlorite prote
41	98.5	5.5	1160	2 T40589	paraportal crystal
42	98.5	5.5	1161	2 E86626	transcription-repa
43	98.5	5.5	1308	2 A90428	serine proteinase,
44	98	5.4	389	2 H90083	polyadenylate-bind
45	98	5.4	429	2 A81315	probable MCP-type

## ALIGNMENTS

## RESULT 1

T47118  
thiamine pyridinylase (EC 2.5.1.2) precursor [validated] - Paenibacillus thiaminolyticus  
N:Alternate names: pyrimidine transferase; thiaminase I  
C:Species: Paenibacillus thiaminolyticus  
C>Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 03-Jun-2002  
C/Accession: T47118  
R/Costello, C.A., Kelleher, N.L., Abe, M., McLaflerty, F.W., Begley, T.P.  
J. Biol. Chem. 271, 3445-3452, 1996  
A>Title: Mechanistic studies on thiaminase I: overexpression and identification of the  
A/Reference number: 224363; PMID:96216437; PMID:8631946  
A/Accession: T47118  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-409 <COS>  
A/Cross-references: EMBL:U17168; NID:9984284; PIDN:AA044156.1; PID:9833806  
A/Experimental source: strain 168  
C/Complex: monomer [validated, PMID:96216437]  
C/Function:  
A/Description: EC 2.5.1.2 [validated, PMID:96216437]  
C/Keywords: transferase  
F/1-30/Domain: signal sequence #status predicted <SIG>  
F/31-409/Product: thiamin pyridinylase #status predicted <MAT>

Query Match 12.8%; Score 230; DB 2; Length 409;  
Best Local Similarity 25.2%; Pred. No. 1.1e-08;

Matches 90; Conservative 53; Mismatches 146; Indels 68; Gaps 14;

QY	7	TLTVGLFPPYLPSPWNENGENEYKILNKDVLPTQVSGNIEYTFEDCYSDASLSQSLPVPFS	66
DB	40	TLKVAIYIPVP-----DPAKFOAAVLDQKQROEPGYKLEFTMDYSADPPDDL-DVFLV	92
QY	67	TDSIFLPYLVSLG-----GVKSLDESLVRCVTGLSHFVSSASVNGSVYGFPOYLCSNF	121
DB	93	LDSIFLPSHFVDAGYLLPFGSDIDQA-----EDVLPLFALQAKRNGEVGLQILICTNL	146
QY	122	LSSPMTQ--QASSLLELAQKVG---YEQIYVPD-----VASSSFTYVGLYQQLQSS	171
DB	147	LFRKQGLKIGQVDNIVELYKKIGTSHSEQIPPOKMGILLINNAAGTTKASWYLE-----	201
QY	172	SSAAVDIKAS-----DPSQSDQVNDITOKKYRTILST-----VVASQREYINSYKQ	220
DB	202	--LIDVTGQYTYDILLPPLDPLNDKVIKRLRLINNAAGEKPSQYVPEDGDAYVRSWFA	259
QY	221	KPISNYVYGSESMCEIKDIIRDQYVNVQIGTS--DKPYVTVDLALNSNLCDKQKV	277
DB	260	QSGGRAFIGVSESMMKMGDYA--EQVRFKRISSSAQODIPLFSDVVSVMKTAHP----	313
QY	278	AVGVIKLTLNTLVLDLGLGL-----TLPAKKGKIALAKSSNFYQQLSQ	323
DB	314	--ELAKKLAVMASADTVEQALRPQADGVPOYLTPARQVVEALMQDPIYVELAQ	368

## RESULT 2

hypothetical protein TRK14.1 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #ext\_change 31-Mar-2001

C/Accession: B96827

R:Theologian, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

ansen, N.F.; Hughes, B.; Hinz, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Mailli, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; M01D:21016719; PMID:11130712

A/Accession: B96827

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1248 <STO>

A/Cross-references: GB:AE005173; NID:g4835752; PIDN:AAD30219.1; GSPDB:GN00141

C:Genetics:

A:Gene: TRK14.1

A:Map position: 1

Query Match 6.4%; Score 116; DB 2; Length 1248;

Best Local Similarity 19.1%; Pred. No. 6.7; Mismatches 123; Indels 126; Gaps 17;

Matches 75; Conservative 69; Mismatches 123; Indels 126; Gaps 17;

14 PYLPSW-----NENGNEVKLIN-----LIKD-VLPQVS-----41

26 PFMFDQGSASAMRPNNSGSDYKAVHNPSIQGEFSLFPMRDYIPQSSNPNGAGD 85

42 -GYNIEYTER-----DCYSDASLOS-----LPDVSTDSIFLPLVLSIGVKS 84

86 MNVTGYMEIRGLIGISHTGSECAVSRSFVENGTSIDERTNSLHEFNKLNHVQSA 145

85 DESLVR--GVTGDHAFSSASAVNGSVYGFPOYLCS--NFLSSPNTQ-----QASSL 135

146 PQALLSKDSSVGNLHGKNTSSASGSVTAKVILCSFGKILPRPGDSKILRYVGGEHI 205

136 LELAKQYBQI-----VPPDVASSSFTVFGLYQ-----QLQSSSSAAVDIKASD- 182

206 ISIKDLSWGLRQKILEIY-----YQTRVVKYQLPGEPLDLVSVSSSED 251

183 ----LPSGDQVNDITQKYRTIILDSIVASQREYINSVKQKPISNYYVGSSEMCERK 238

252 LQNNLEENEMENRGSGKIMPLFS--ISDMDDLGVNKNDDGSEF-----297

239 DIIRDQYNVOLIGTSKPRYYTVDVLANSLCEBKQVAVKINLITNTLVLDLGLG 298

298 ----GYVAVNMWDIGSGKNSITLLGLDSSANNLAEILDVNRTEGI--NTIAGDVVG 349

299 LILPANKGIAHLAKSSNFYAQLSQDFPKASE 331

350 -----ASGLMNGFQQTSAQOSE 367

## RESULT 3

phycoobilisome linker protein apce - red alga (Porphyra purpurea) chloroplast

C:Species: chloroplast Porphyra purpurea

C>Date: 19-Mar-1997 #sequence\_revision 09-May-1997 #ext\_change 10-Sep-1997

C/Accession: S73184

R:Reith, M.; Munholland, J.

Plant Mol. Biol. Rep. 13, 333-335, 1995

A>Title: Complete nucleotide sequence of the Porphyra purpurea chloroplast genome.

A:Reference number: S73108

A/Accession: S73184

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-886 <RET>

A/Cross-references: EMBL:U38604; NID:g1276652; PID:g1276729

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995

C:Genetics:

A:Gene: apce

A:Genome: chloroplast

C:Keywords: chloroplast

Query Match 6.4%; Score 115.5; DB 2; Length 886;

Best Local Similarity 20.5%; Pred. No. 4.3;

Matches 75; Conservative 57; Mismatches 124; Indels 109; Gaps 17;

54 SDASLQSLPVSSTDSIFLPLVSLG-GVKSLSDELVRGVTGL-----HSFVS--SS 103

521 SSANFNSVDVIAKVS-----YLRVFGREYQGEKLLRPISQLQDQGISVREFIRQLAK 576

104 ASVNGSVYGFPOYLCS-----NFLSSPN-GTQOASLLELAKQYEQIYVPDVASS 156

577 SSIFRSLYWEPLTYCAIEYHNRLGRPYGKQENKRFDIYKGYOVVDAILDSPE 636

157 SFTVFG-----LYOQLQSSSSAAVDI-KASDLPOSGD 188

637 YITFGENVVPRERYTTPAGIALRSIRPGIIDRFKVKVSSKSRFVGLGKVEIRSSND 696

189 -----QNKDITQKRTIILDSIVASQREY--INSVQKQKPISN 225

697 IQSRIAGVTALRQGVIFPDVNQSSQE--VLEQALRAVROIFERDLNFSIGEFLLD 753

226 YVYGSSSMCEIIDIRD-----QQY--NQQLISTQK-----PYVYTDVLAL 266

754 TESSFLNKQIVNKLQKALSELVYKREYQYPNPKVIELGTHKILGRAPNNQAEIRFL 813

267 NSNLCEKQKVAEIVKINLL-----TNTLVLDLGLGLTLPANKGIAHLAKSSNFYA 319

814 NQLASGLSTFETLVNNSSEYDSVYGTNVVPRRPP--TLPA-----ANFPTETLVN 865

320 QLSQ 324

866 RLTKQ 870

## RESULT 4

hypothetical protein SSO1375 [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C>Date: 24-May-2001 #sequence\_revision 24-May-2001 #ext\_change 24-May-2001

C/Accession: B90294

R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder,

arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A99139

A/Accession: B90294

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-613 <XUR>

A/Cross-references: GB:AE006641; NID:g13814587; PIDN:AAK1609.1; GSPDB:GN00155

C:Genetics:

A:Gene: SSO1375

Query Match 6.2%; Score 111.5; DB 2; Length 613;

Best Local Similarity 19.5%; Pred. No. 4.7; Mismatches 113; Indels 165; Gaps 22;

Matches 85; Conservative 72; Mismatches 113; Indels 165; Gaps 22;

13 FPY---LPSNENG-----NEVKLINL-IKDVLPIVQSGYN-----44

88 YPVNQVWVNPENGLPLNMSYDNMSIHEISFIALPTVDVYPTKXI--YNSSLYLGKGLLS 145

45 -----LEYEPFCYSDASLQSLPD--VFSTDSIFLPLVLSLGSV-----KSIDES 87

146 ATYNNIAIIISYKDFSKY-NTLSQYPSHILQGTGIGL-FIIPLOALPWVYKSVNS 203

QY 88 LVR-----GVTGDLHSFVSSASVN-GSVYGFPPQYLCNPL 122  
 Db 204 VIOQTYIFQNETNLLINGNDFGSLTQLGLYDELSTVQTLFEPLNTTSPY-----SSIL 256  
 QY 123 LSPNGTQOA-----SSLELA-----QKVGVEQIV----- 148  
 Db 257 IDSINNTEAMANNRPITDINQITDINSIIDLAYSSSTLTPGEQEFAYIVSFGNITTKIES 316  
 QY 149 -----YPDVASSSSFTVFGLYOQLLOSSSSAANDIKASDLPQSGDQVNDKI 194  
 Db 317 EIVSELPLVNLVIOYPSISNOSPITIVSFVNLNRSATKVEGVVGLLNYSTG----- 369  
 QY 195 TOKYRTIDSTVVASQREYINSVKQKPISNVY-----VGYSESMCEIKIIRD--QOY 246  
 Db 370 -----LISKISIMNSNGELVFNVSPGTIVLYVYHYPNLGLNTEYWGQIRIDVRSGINKY 423  
 QY 247 NVQLIGTSKRYVYTDVLAALNSLNCDEKQKAVEYIKULNTLVLDLGLGLTIPANKN 306  
 Db 424 NF-----TRVEPMYINII-----DTSIDSGFKISVKLIINPL--NQTVSGQLYIMIS-----N 468  
 C 307 GIAHLAKSNFYAOL 321  
 Db 469 GIASSASAPTESOL 483

RESULT 5

T37919  
 GCN1 homolog - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 07-Dec-1999  
 C:Accession: T37919  
 R:Conor, R.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.; Wood, V.  
 submitted to the EMBL Data Library, December 1995  
 A:Reference number: Z21754  
 A:Accession: T37919  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-2670 <CON>  
 A:Cross-references: EMBL:Z68198; PIDN:CAA92385.1; GSPDB:GN00066; SPDB:SPAC1866.05C  
 A:Experimental source: strain 972h-; comid c1866  
 A:Genetic8:  
 A:Gene: SPDB:SPAC1866.05C  
 A:Map position: 1  
 A:Introns: 50/1

Query Match 6.2%; Score 111.5; DB 2; Length 2670;

Best Local Similarity 19.7%; Pred. No. 45;  
 Matches 77; Conservative 60; Mismatches 146; Indels 107; Gaps 17;

C 10 VGLF--PYLP-----SMNNGNEVKLINL-----IKDVLPTQVSGYNI 45  
 Db 1433 LGIYFEPYLPDLLPLLLTSLFGDNANEVREATMDAVKQIMSQLSAGVALLPTLLDGLN- 1491  
 QY 46 EYTFEDCYSDASLOSPLPVSTFD--SIFLPYLVSLGVKSIDESLVRGVTDGLHSFVSS 103  
 Db 1492 EYNNRSKASAVEILGLMAYMPKQLSVLPFI-----IPKLSEVLT-----DSHQVANT 1541  
 QY 104 ASVN-----GSVYGF--OYLCNPLSSPNGTQOASSILELAQKVGQIYIPV- 152  
 Db 1542 ANKSLRFGDVISNPEIQTLVPLTKALSDCTRYTDALTEALLKTSFVHYLDPPSLALVI 1601  
 QY 153 -----ASSSFTVFGLYOQLLOSSSSAAV-----DIKASD 182  
 Db 1602 PILKYLGRNAGTKRGQAKIFGLMASTLTPENLAVYIESLMPRLREVLIDPVPDTRATA 1661  
 QY 183 LPQSGDQVNDKITOKYRTILDS--TVVASQREYINSVKQKPISNVYVGYSESMCE--IK 238  
 Db 1662 AKALGSLIEKGEKKFPLIEPLFVNLRSCESEVDRQGAQGLSITLGLGLARLEVDLP 1721  
 QY 239 DIIRDQ-----YVQLIGTSKRYVYTDVLAALNSLNCDEK---QKAVE 280  
 Db 1722 EILKNTSSPVPHIREPSISLIYLPATGSRFPQYLABAIPILISGLADDELVOTASLR 1781

QY 281 VIKULLTN--TLVLDLGLGLTIPANKNGI 308  
 Db 1782 AAKMIVANNVATKSVDLL-----LPELEKGL 1806

RESULT 6

T06799  
 fimbrin/plastin-like protein - wheat (fragment)  
 C:Species: Triticum aestivum (common wheat)  
 C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 11-May-2000  
 C:Accession: T06799  
 R:Crüz-Ortega, R.; Cushman, J.C.; Ombry, J.D.  
 Plant Physiol. 114, 1453-1460, 1997  
 A:Title: cDNA clones encoding 1,3-beta-glucanase and a fimbrin-like cytoskeletal protei  
 A:Reference number: Z15824; MUID:97422889; PMID:9276954  
 A:Accession: T06799  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-448 <CRU>  
 A:Cross-references: EMBL:U67717; NID:G1575594; PIDN:AA049813.1; PID:G1575595  
 A:Experimental source: cv. Victory; root; 4 days old  
 C:Function:  
 A:Description: actin-binding  
 C:Keywords: actin binding; alpha-actinin actin-binding domain homology; calmodulin repeat

Query Match 6.0%; Score 109; DB 2; Length 448;

Best Local Similarity 20.7%; Pred. No. 4.4; 155; Indels 68; Gaps 17;  
 Matches 77; Conservative 72; Mismatches 17;

QY 16 LPSNNGNEVKLINLIKDV-----LPTQVSGYNIETFPDCYSDASLQS 60  
 Db 5 LNPERNENHTLCLNSAKAIGCTVVGIGFVGRHVLVGLISQIKIQLADLNLKK 64  
 QY 61 LPDVFS--TDSIFLPYLVSLGVKSLDESIVR-----GVTGDLHSFVSSASVNGS 109  
 Db 65 TPQVLELPDSDKIDDEVLT-----STEKMLRWNNHHLKKAQYKKTNNF--SSDVADGE 118  
 QY 110 VYGF-----POYLCNPLSSPNGTQOASSILELAQKVGVEQIVP-DVASSSFTVFG 162  
 Db 119 AYAILLAALAE-TSPETLETKNPDERAKNVLEQAELOCKKRLTPKDIPEGSSANLNA 177  
 QY 163 LYQQLQSSSSAAVDIKASDLPQSGDQVNDKITOK---YRTILDSTVVASQREYINSVKQ 219  
 Db 178 FVAQIFQHRNGLTSDIQVLTQSASRDVLSSEERAFPMWINSLEVES---YLNNVFE 234  
 QY 220 GKPISNVYVGYSESMCEIKIIRDQVNVOLIGTSKPYVYTDVLAALNSLNCDEKQYAV 279  
 Db 235 --DVNRGMV-----LLEVLDKVPSPGVNMWL--ESKPKIKLPKRLKN--CNOVVIGK 282  
 QY 280 EV---INKULNTLVL--DLGLGLTIPANKNGIAHLAKSNPFYAOUSQOFDAKESEVRV 334  
 Db 283 ELKPSLVNLAGNDIVQGNKKLIVALLQMLRFNLQLNLRSHRSKSGSQGQITDADI-- 340  
 QY 335 LRCVDFANKEVK 346  
 Db 341 ---LNMANSKVK 349

RESULT 7

T32787  
 hypothetical protein F14D2.6 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: T32787  
 R:Du, Z.; Le, T.T.  
 submitted to the EMBL Data Library, December 1997  
 A:Description: The sequence of C. elegans cosmid F14D2.  
 A:Reference number: Z21225  
 A:Accession: T32787  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-998 <DUZ>

A:Cross-references: EMBL:AF040643; PIDN:AB994960.1; GSPB:GN0002; CESP:FlAD2.6  
A:Experimental source: strain Bristol M2; clone FlAD2  
C:Genetics:  
A:Gene: CESP:FlAD2.6  
A:Map position: 2  
A:Introns: 50/1, 128/1, 189/3, 210/2, 277/3, 391/1, 413/3, 460/1, 499/1, 577/1, 641/1, 7

Query Match	6.0%:	Score 108.5:	DB 2,	Length 998;
Best Local Similarity	22.6%:	Pred. No. 16;		
Matches	76;	Conservative	56;	Mismatches 125;
				Indels 79;
				Gaps 17;

Qy 8 LTVGLEPFLPSMNENGNKVLINLIK-DY-LPTQVSGNIEYTEFPCYSDA-----SLQ 59  
:::|::|||:::|:::|:::|:  
Db 137 LLTSIYNISIDWWSQETSKWDINNTKLPGVICRVAERDI-YADFAAYGNRKVVAAADSYK 195

QY 60 SLPDVFSTDSIFLPYLVLGGVKSLSDELVRGVG-DLHSFVSSASVNGSVYGFPG-YL 117  
 Db 196 STTQGFPTIGTALQVLSLSLKKLNGMDEIYETGLENVSFIGNLETHNGKNGGLPEKYM 255

118 CSNF--LLSSNGTQQASSLLELAKVCGEQIVPDVASSSFTV-----FGLYQOL 1674  
256 SFNFPSIHNDEN-----LRITGDSL--KQLPGAGKFTIHIANNHPDFCLSTSE 3022

QY 168 LQSSSAVDLKASDLPQSGDYNKIDITQ-----YRTILDSTVAS 209  
 || : | : || : | :  
 Db 303 LQVFARTNLKIVGEV-----QICKDLFRKDGQKTCIFKLSMLBPKCQHILGDIPINAT 357

QY 210 QREYINSVKOGKPRISNY-YVGYSESMCEIKD-----IIRDOQ----YNVOLGTSBKPY 258

Db 358 NEEYLINLKMTKVICGFEGFGAGACSCFGERYGDIPILIRSSNKLRLRVTLTPQMKNPPE 417

```
QY 259 VYTD--VLANSN-----LCDEKQKVAVEVIK 283
      | : | : | : | | | : ||
Db 418 SPTNNPVIWNNTEQIFQTRACQEFQKFTKSIK 453
```

RESULT 8  
T22583

C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000  
C:Accession: T22583  
B:Barlow: Y

submitted to the EMBL Data Library, July 1996  
A:Reference number: Z19584  
A:Accession: T22583  
A:Status: provisional; translated from CD/MDM/MDMT  
A:Status: provisional; translated from CD/MDM/MDMT

A: molecule type: DNA  
 M: residues: 1-528 <WIL>  
 A: cross-references: EMBL:277663; PIDN:CAB01214.1; GSPDB:GN00023; CESP:F53B  
 A: (Protein) source: *Escherichia coli*

C:Genetics:  
A:Gene: CESP:F53F4.5  
A:Map position: 5  
A:Introns: 47/3, 110/3, 151/1, 109/3, 353/3

C;Superfamily: dimethylaniline monooxygenase (N-oxide-forming)

Matches 59; Conservative 47; Mismatches 80; Indels 86; Gaps

QY 43 YNIEYTFDCYSDASLQSLPDVFSTDSILPLYL-----SLGVK----- 82

Db 238 YDVQL--FSRYYDTLTKTI PHAVAND--PMEYRLQQRMDHDVYGLNPDHRRFQQHPYVND 239  
 Qy 83 SLDESLVRG---VTGDIHSFVSSASAVNGS-----VYGFPPQYLCNFTLSS 125

Db 294 ALANLLCAGYITITBEDIDFTENSIVKGRFKCDIFLCTGYTGFP-----FV---- 344

Db 345 -----DSDIVEIKNQVPLYKTVFP--ENSDSVAVIGLLPIGSIAPISEIQRMARV 396

```

Oy 177 DIKASDLPQSDQDNKXITQCYRFTILSDTVAASREKINSVKOSKPISNYYVGSSEMC 236
Db 397 FAGRCQPLPSSQEQID-DIQKK-----KAAMKKRYFDSIKH-TIQVDYMSYDEIAE 445
Oy 237 IKDIIRDQYN-----VGLISTGSKPYVY 260
Db 446 IIGCLPEMKHILFFYERFEMKPLFGAANVPIAY 477

```

## RESULT 9

C:Species: chloroplast *Aglaethamnion neglectum*  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 06-Feb-1995  
C:Accession: S20621; S61554

submitted to the EMBL Data Library, April 1992  
A:Description: Characterization and transcript analysis of the major phycocyanin gene in *Chlorella*  
A:Reference number: S20621

A;Molecule type: DNA  
A;Residues: 1-885 <APT>  
A;Cross-References: EMBL:Z11905; NID:g22706; PID:g22707

A:Gene: apcE  
A:Genome: chloroplast  
C:Keywords: chloroplast

Query Match	6.0%;	Score 108;	DB 2;	Length 885;
Best Local Similarity	21.2%;	Pred. No. 15;		
Matches 79;	Conservative 70;	Mismatches 119;	Indels 104;	Gaps

QY 53 YSDASLQSLDPVFS-IDSIPLP-----YLVLGGVKSLDESLVRGVTGDLHLSFVSS--S 103

QY 104 ASVNGSVYGEPPQYLCSNF-----LSSPN-GQQASSLELAQRKXGEQIVPVASSS 156  
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Db 577 SSVERSLYWEPPLYICKAIEYTHRLGLGPPSYGROEINQIFDIVREGYYKMIDYLLNSSE 636

```

QY 157 SFTYFG-----LYQQLLQSSSSAAVDIKASDL--PQSGDQV 190
      ||      ||      ||      ||      ||      ||      ||
Db 637 YIRSFQDNIVPYERYITSANMILKSNSNLFYSTLKTQNSNSKFIELSNIIEKRLSNI 696

```

```

QY 191 NKDITQKYRTLDS-----TWASQ---REYINSVKQ---GKPISNYYVCSSESCB-- 236
      ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 697 QARISQGVSTVRDQYKVFQLTQLSSQVDKQVVKAIYRQIFERDINSFATGDEFINLEKA 756

```

```

QY 237 -IKDIIRDQYNNQLGTSD-----KPYVVDLALN-----SNLCDEK-----QK 216
      : : | | : : : | | : | : |
Db 757 LVNNNITVQGF-IEQIGSSSLGREGFYQYPNTKVIETLGTKHFLLGRAPNNQAEIRIYYNQI 815

```

```

QY 277 VAVEIKNLLNTLTV-----DLGLGL-----TPANKGIAHLAKSSNFYQJLS 322
      ::::: || ::::: ||| | : | | :
Db 816 LASGLISFIT--VLVNSNEYNQVFGVNIIVPYRRFTTLP-----ANFPTKLYNTLT 867

```

```
QY      323 QQFPAKESVRYV 334
      :|      ||:|
Db      868 KQ---SSEIIV 875
```

RESULT 10  
A96714  
bimathematical notation MCR 13 (improved) - hybridoma thalians

C;species: Arabidopsis thaliana (mouse-ear cress)  
C:date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C;accession: A96714  
C;description: A cDNA from Drosophila melanogaster, encoding a protein of 108 amino acids.

Chun, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.; Jensen, N.F.; Hughes, B.; Huizar, L. Nature 408, 816-820, 2000

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maltz,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: A96714  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1033 <STO>  
A:Cross-references: GB:AE005173; NID:g6358765; PIDN:AAF07347.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: T6L1.12  
A:Map position: 1

Query Match 5.9%; Score 106.5; DB 2; Length 1033;  
Best Local Similarity 20.6%; Pred. No. 24;  
Matches 96; Conservative 73; Mismatches 135; Indels 163; Gaps 26;

QY 1 MSTGPKTLTVGLFPYPSWNGNEVKLIN-----LIDVLPTQVSG 42  
D 85 LSNMGNATIG-----DGSKEINNGQOKISAKHTPDIVSEGMEDLPTDP-- 132  
QY 43 YNITTEPDCSDASLQ--SLPDV--FSTDSIFLPLVLS--LGVKSLDESIVRGVTGD 95  
Db 133 -EFSYESYMYSESOSQWTDIPDIPSKSTD-----VSROKIGNHSESOSLVTEIP-D 183  
QY 96 L-----HSFVSSASVNGSVGFPOYLCSNFLSPNGTQO-----ASSLLE 137  
Db 184 IPSGSTVSSQKRIKGNLSKSSQSTETIPDIPS-----OSSNASSQKRYGNLSSESLSM 235  
QY 138 LAQ-----KVGVEQIVYPDVAVSSSFTVFGLYQQLLOSSSAVAIDIKAS-----DLP 184  
Db 236 LPQVTOGMEPPYQAFIC-----LTKEIMEDPVTTTCGTCERQAVIEWPDSF 283  
QY 185 QSGQOVNKDIT-QYRTITLDSVVASQREYNYSVKQKPISN-----YVVGYS 232  
Db 284 GNSDEINCPVYGQKLTTELISANVV-----LKTIIQEWKVENAARIKVAHAALSLIGSE 337  
QY 233 SMCEIKIIRP-----QQVN-----VOLIGTSDKRYVVT-----DVLALNSNL 270  
Db 338 SM--VIDLRDLQMTCEGKEYNKVQVREAGITQLL--DRYLTVRSKQVRELLKFLRTL 392  
QY 271 CDEKQKAVAEVYKMLTFTVLVDLIG-----LGLTLPANKN-----GIALHAK 313  
Db 393 ADEETDQKENVITITMTCYIKLLGSSHQVRHAQKLLLESLSGACKEKIGARAI 452  
QY 314 SSNRYAQLSQFD--AKESEVRVLRCDVPANKEVNCA--GVLRPF 356  
D 453 LMLVTAKYNRELDSFASSETSDOILRNLEKCPENIKOMAESGLLEPLL 499

RESULT 11  
A90551  
conserved hypothetical protein MYPU\_3130 [imported] - Mycoplasma pulmonis (strain UAB CT  
C:Species: Mycoplasma pulmonis  
C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001  
C:Accession: A90551  
R:Chamand, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;  
Nucleic Acids Res. 29, 2145-2153, 2001  
A:Title: The complete genome sequence of the murine respiratory pathogen *Mycoplasma pulm*  
A:Reference number: A95512; MUID:21267165; PMID:11353084  
A:Accession: A90551  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-2819 <KUR>  
A:Cross-references: GB:AL445566; PID:g14089727; PIDN:CAC13486.1; GSPDB:GN00153  
A:Experimental source: strain UAB CTIP  
C:Genetics:  
A:Gene: MYPU\_3130  
A:Genetic code: SGCG

Query Match 5.9%; Score 106.5; DB 2; Length 2819;  
Best Local Similarity 20.8%; Pred. No. 1.1e+02;

Matches 77; Conservative 66; Mismatches 105; Indels 123; Gaps 19;

QY 25 EVKLINLIKDVLPQVSGYNIETFPDCYSDASLQSLPDVSTDSIFLPLVLSGVKSL 84  
Db 1358 EIKSSNIIOKRLGGTITALVTNCE-KIPSDTELQSLSTLIN-----ESL 1401  
QY 85 DESLVRGVTGDLHSFVSSASVNGSVGFPOYLCSNFLSPNGTQOASSLLEAOKVY 144  
Db 1402 DLFKTNINISNLVDLVASALIN-----QSSSVL---TKV-- 1434  
QY 145 EQIVYPDVAVSSSFTVFGLYQQLLOSSSAVAIDIKASDLQSGQVKNKDTOKRTILDS 204  
Db 1435 ENITHKALDOQNSY--PAVITKTL-----SVKEDVDVKSD--FKAIKKKFLNLF-- 1480  
QY 205 TVVASQREYNYSVKQKPISNYVGY-----SESCEIKIIRPQO---YVVOIGT--- 253  
Db 1481 ---AKNDYL-LTKFISPIVDKTIIDFDSATKSSIVKLKIVLSQBLKRSQIITDITFA 1535  
QY 254 -SDKPYVY-----TDVLALNSNL-----CDEKQKAVE--VINKLTL--- 287  
Db 1536 NKNKYLAIVENLDFLIDALVDNKENIINLSNTLGKIKTDEPNKIIFITKXVLMTSDS 1595  
QY 288 -----NTVLVDLGLTLPAKNGIAHLAASSNRYAQLSQQFDA-----KESEVRV 334  
Db 1596 SITEEINTTIV-----KSAHELISIVDSNLFNLQKLFALLERKMKDSKSTS 1644  
QY 335 LRCVDFANKV 345  
Db 1645 LEFSTFMKSI 1655

RESULT 12  
D75208  
sugar abc transporter binding protein related PAB2439 - Pyrococcus abyssi (strain Orsay  
C:Species: Pyrococcus abyssi  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
C:Accession: D75208  
R:anonymous, Genoscope  
submitted to the EMBL Data Library, July 1999  
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome str  
A:Reference number: A75001  
A:Accession: D75208  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-545 <RAM>  
A:Cross-references: GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CAB49115.1; PID:g54576-  
A:Experimental source: strain Orsay  
C:Genetics:  
A:Gene: PAB2439

Query Match 5.9%; Score 106; DB 2; Length 545;  
Best Local Similarity 25.0%; Pred. No. 9.7;  
Matches 38; Conservative 24; Mismatches 68; Indels 22; Gaps 6;

QY 19 WNEGNEVKLINLIKDVLPQVSGYNIETFPDCYSDASLQSLPDVSTDSIFLPLVSL 78  
Db 92 WKENGIVNKIV-IPEDMIRYDQS-FKDQYQEF--LSKQPLQOQADDFVNSGAFLEPNLAEE 147  
QY 79 GGVVSLDE--SLVRGVTGDLHSFVSSASVNGSVGFPOYLCSNFLSPNGTQOASSL 136  
Db 148 GYIIDITDYAKAVGSVNDPFPSLLEAKYKGRLYGFPQ-----DREARPLY 194  
QY 137 ---ELAQRVGEQIVYPDVAVSSSFTVFGLYQ 165  
Db 195 IRKDVACVGLDVSTLPDKVNGEFTWSDVE 226

RESULT 13  
G82934  
hypothetical protein U0094 [imported] - Ureaplasma urealyticum  
C:Species: Ureaplasma urealyticum  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: G82934

R,Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Casseil, G.H.  
submitted to GenBank, February 2000  
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mit  
A:Reference number: A82870  
A:Accession: G82934  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1054 <GLA>  
A:Cross-references: GB:AE002109; GB:AE222894; NID:G6899045; PIDN:AAF30500.1; GSPDB:GN001  
A:Experimental source: serovar 3; biovar 1  
C:Genetics:  
A:Gene: UU094  
A:Genetic code: SGC3

Query Match 5.9%; Score 106; DB 2; Length 1054;  
Best Local Similarity 20.2%; Pred. No. 26;  
Matches 70; Conservative 51; Mismatches 110; Indels 116; Gaps 17;

24 NEVKLIN--LIKDVLPQVSGYNIETEPDCYSDASIQSLPQVSTDSIFLPIVLISLG 80  
372 NELKLNKQILIKDIEVAFSTGY-----DVAIDISGISP----- 405  
81 VKSLDESLVRYGTGLH--SFVSSASVNGSVGFPQYLCSNPLSSPNCQOASSLLE 137  
406 ----NLLKGVANDLNKKEFLQSSQDFRMLTYNITQ-----QNNLTMILITLIT 449  
138 LAQKVG-----YEOIVY--PQVASSSFVFGVLQQLQSSSAVADIKASDLPQS 186  
450 SNQSIGMWIDGPKPLKELIYNPSIDQSVKTIW-----SKTSMDFKNNLIMEI 500  
187 GDQVNDITQKRYTILDS--TVVASQREYINSV-----KQKQPSINYVYGSSSM 234  
501 G-----QIRTLDSISLIEIKTYKPIINKLELISLIEARSKQPKNY--GFIDGL 549  
235 CEI-----KDI-----IRQOQNVOLIGTSDPRYTTDVLAISNLCEKQKVA 278  
550 DELSFSPFLNKQNDLPSSINFKIGQNTLYDLITEVK--IANNLIPHENQYDPKDOI 607  
279 VEVINKLTLNTLVLDLGLGLTPANKNGIAHLAKSNFYAQLSQOF 325  
608 IST-NYPLTKIKILDLSL-----NKNDQLDYHKGVEMTSLKKF 647

RESULT 14  
AB1746  
hypochemical protein lin2511 [imported] - Listeria innocua (strain Clp11262)  
C:Species: Listeria innocua  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
A:Accession: AB1746  
A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madheno, E.; Maitournam, A.; Ma  
ker, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baguerio, F.; Berche, P.; Bloeker  
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Feihl, H.  
D.; Jones, L.M.; Karet, U.  
Science 294, 849-852, 2001  
A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AB1746  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-358 <GLA>  
A:Cross-references: GB:AL592022; PIDN:CAC97738.1; PID:G16415033; GSPDB:GN00178  
A:Experimental source: strain Clp11262  
C:Genetics:  
A:Gene: lin2511

Query Match 5.9%; Score 105.5; DB 2; Length 358;  
Best Local Similarity 20.9%; Pred. No. 5.5;  
Matches 62; Conservative 44; Mismatches 90; Indels 101; Gaps 13;

125 SPNGTQOASSLLELAQ--KVGVEQIVPQVASSSFVFGLYQO-----LLQSSS 172  
53 SDNYIKKLGSIDLDNLQKKSNQYKISVDSSSSYSASVLMKKEPLDLTLHALQNSE 112

173 SAAVDIASDLPQSGDV-----NKDITQKRTI-----LD 203  
113 DGRAYIPVADIDYADADSFSSLLSDSAAQIFNKVLAKQKDESKYLANFETFOFNSNOTID 172  
204 GTVASQREYINSVKQKQPSIN- YVGYSE-----SMCEIKDIIRPOQY 246  
173 SETYDRQAEELKIEASGAIYSYLNNDLDDKFTSKDNGDIVLNLKNEIKDVLND-- 229  
247 NVQLIGTSKRYVYT-----DVLAISNLCEKQKVAE-----VINKLTLNT 289  
230 -----VFTALDNGNITLTLSEIDGSSQKRAQKKMNGEKSYSALKNLNSK 277  
290 LVLDLGLGLTL--PANKNGIAHLAKSNFYAQLSQOFDAKSE--VRLRCVDFPANK 344  
278 --NQLTFKLTLPDSKSKGFSKAVITTNFE-----DSKTKDILNTTIVDMIDYE 325

RESULT 15  
E70224  
hypochemical protein BBE02 - Lyme disease spirochete plasmid E/1p25  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 08-Oct-1999  
A:Accession: E70224  
A:Authors: Smith, H.O.; Venter, J.C.  
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.  
A:Reference number: A70100; MUID:98065943; PMID:9403685  
A:Accession: E70224  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1277 <KLB>  
A:Cross-references: GB:AE000785; NID:G2689951; PIDN:AA66031.1; PID:G2689960; TIGR:BBE0  
A:Experimental source: strain B31  
C:Genetics:  
A:Genome: plasmid

Query Match 5.9%; Score 105.5; DB 2; Length 1277;  
Best Local Similarity 22.1%; Pred. No. 38;  
Matches 76; Conservative 56; Mismatches 107; Indels 105; Gaps 21;

15 YLPSMNENGEVKLINLIK-DVLPQVSGYNIET-----FDCYSPASL 58  
784 HENYRKLG--IKQNIKQELISKDHPUSIEENKNSIATANNYKFPDKFTSGG- 840  
59 QSLPVESTDSIFLPYVLSIGVKSLSDESIVRGYGLHSFVSSASVNGS-----VYG 112  
841 -----DENLFRYVTF-NKLKIK-----GNL-TLVPSALWNESSRILRKHI 883  
113 FPQYLCNPLISSPNGTQOASSLLELAQKVGEOIVPQVASSSFVFGLYQQLQSSS 172  
884 FARYKL-NYIVQFNKKR-----FQVHSSFRFAIFOL-SNIKESTS 923  
173 SAAVDIASDLPQSGQVNDITQKRTIILDSITVASQREYINSVKQKQPSINYVGYSE 232  
924 S-----FRAKEMIQSSDNLKKEITDLDKSDKDAYGIELN-INQIKKLSPI-----QE 971  
233 SMCEIKDIIRPOQYNAVOLIGTSKRYVYTDLAISNLCEKQKVA--AVEYINKLTLNTL 290  
972 SIIFKQD--NEFT--LINK-----MFSKFSALGEGVIFPKKLDLSIKRKSLAKE-- 1019  
291 VLDDLGLGLTLPANKNGIAHLAKSNFYAQLSQOF-----DAKES 330  
1020 -----CNKKVLIFLYSGANIH-QFNRSFPEDKAKES 1050

Search completed: January 24, 2003, 19:43:11  
Job time : 28 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2003, 19:41:50 ; Search time 12 Seconds  
(without alignments)  
598.630 Million cell updates/sec

Title: US-09-675-509-4  
Perfect score: 1802  
Sequence: 1 MSTQPKTLTVGLPFLPSMN.....CVDFAKVKNCAGVLRLPL 356

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 122226 seges, 20178551 residues  
T number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA\*  
1: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB\_pdp:\*  
2: /cgn2\_6/ptodata/2/pubppaa/PCT\_NEW\_PUB\_pdp:\*  
3: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB\_pdp:\*  
4: /cgn2\_6/ptodata/2/pubppaa/US06\_PUBCOMB\_pdp:\*  
5: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB\_pdp:\*  
6: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB\_pdp:\*  
7: /cgn2\_6/ptodata/2/pubppaa/PCTUS\_PUBCOMB\_pdp:\*  
8: /cgn2\_6/ptodata/2/pubppaa/US08\_PUBCOMB\_pdp:\*  
9: /cgn2\_6/ptodata/2/pubppaa/US09\_NEW\_PUB\_pdp:\*  
10: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB\_pdp:\*  
11: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB\_pdp:\*  
12: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB\_pdp:\*  
13: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB\_pdp:\*  
14: /cgn2\_6/ptodata/2/pubppaa/US60\_PUBCOMB\_pdp:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100.5	5.6	745	8 US-08-834-666A-6	Sequence 6, Appl
2	97.5	5.4	2037	9 US-09-951-402-3	Sequence 3, Appl
3	97.5	5.4	2037	10 US-09-951-401-3	Sequence 3, Appl
4	97.5	5.4	2037	10 US-09-922-101-3	Sequence 3, Appl
5	97	5.4	615	9 US-09-992-598-194	Sequence 194, App
6	97	5.4	615	9 US-09-989-293A-194	Sequence 194, App
7	97	5.4	615	9 US-09-989-735-194	Sequence 194, App
8	97	5.4	615	9 US-09-990-444-194	Sequence 194, App
9	97	5.4	615	9 US-09-989-730-194	Sequence 194, App
10	97	5.4	615	9 US-09-989-436-194	Sequence 194, App
11	97	5.4	615	9 US-09-991-181-194	Sequence 194, App
12	97	5.4	615	9 US-09-993-687-194	Sequence 194, App
13	97	5.4	615	9 US-09-989-734-194	Sequence 194, App
14	97	5.4	615	9 US-09-997-653-194	Sequence 194, App
15	97	5.4	615	10 US-09-989-722-194	Sequence 194, App
16	97	5.4	615	10 US-09-989-723-194	Sequence 194, App
17	97	5.4	615	10 US-09-989-279-194	Sequence 194, App
18	97	5.4	615	10 US-09-989-727-194	Sequence 194, App
19	97	5.4	615	10 US-09-989-731-194	Sequence 194, App

20	97	5.4	615	10 US-09-989-732-194	Sequence 194, App
21	97	5.4	615	10 US-09-991-073-194	Sequence 194, App
22	97	5.4	615	10 US-09-990-442-194	Sequence 194, App
23	97	5.4	615	10 US-09-991-163-194	Sequence 194, App
24	97	5.4	615	10 US-09-993-604-194	Sequence 194, App
25	97	5.4	615	10 US-09-990-456-194	Sequence 194, App
26	97	5.4	615	10 US-09-989-721-194	Sequence 194, App
27	94.5	5.2	652	8 US-08-834-666A-16	Sequence 16, Appl
28	93.5	5.2	418	10 US-09-925-302-565	Sequence 565, App
29	93.5	5.2	991	10 US-09-815-242-5803	Sequence 5803, App
30	93	5.2	346	10 US-09-741-669-380	Sequence 380, App
31	92.5	5.1	592	9 US-10-027-806-80	Sequence 80, Appl
32	92.5	5.1	592	9 US-10-034-623-80	Sequence 80, Appl
33	92.5	5.1	607	9 US-10-072-436-11	Sequence 11, Appl
34	92.5	5.1	816	10 US-09-815-242-11849	Sequence 11849, A
35	92	5.1	468	10 US-09-765-272-10	Sequence 10, Appl
36	91.5	5.1	376	10 US-09-801-368-206	Sequence 206, Appl
37	91.5	5.1	722	10 US-09-815-242-12888	Sequence 12888, A
38	91	5.0	424	9 US-09-738-626-4306	Sequence 4306, Ap
39	90	5.0	2437	10 US-09-815-242-5834	Sequence 5834, Ap
40	90	5.0	6281	10 US-09-815-242-12996	Sequence 12996, A
41	89.5	5.0	384	9 US-09-738-626-4937	Sequence 4937, Ap
42	89.5	5.0	886	10 US-09-801-368-180	Sequence 180, App
43	89.5	5.0	963	9 US-09-764-868-877	Sequence 877, App
44	88.5	4.9	4563	9 US-09-870-759-128	Sequence 128, App
45	88	4.9	511	10 US-09-815-242-13865	Sequence 13865, A

# ALIGNMENTS

RESULT 1  
US-08-834-666A-6  
; Sequence 6, Application US/08834666A  
; Patent No. US20020044949A1  
; GENERAL INFORMATION:  
; APPLICANT: Kleantous, Harold  
; APPLICANT: Lissolo, Ling  
; APPLICANT: Tomb, Jean-Francois  
; APPLICANT: Miller, Charles  
; APPLICANT: Al-Garawi, Amal  
; TITLE OF INVENTION: 76 Kda Helicobacter Polypeptides and  
; TITLE OF INVENTION: Corresponding Polynucleotide Molecules  
; NUMBER OF SEQUENCES: 64  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Clark & Elbing LLP  
; STREET: 176 Federal Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/834,666A  
; FILING DATE: 01-APR-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,175  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-428-0200  
; TELEFAX: 617-428-7045  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 745 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single

```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: signal sequence
; LOCATION: 1...20
; OTHER INFORMATION:
US-08-834-666A-6
```

```

Query Match          5.6%; Score 100.5; DB 8; Length 745;
Best Local Similarity 21.1%; Pred. No. 1.6;
Matches 63; Conservative 43; Mismatches 112; Indels 81; Gaps 11;
```

```

QY 103 SAVNGSYGFGPPQYLCSNFFLLSPNG----- 128
D 212 TADINGGVYQCK--AKNGSSSSNGSGSSTQTATTGTGGVITTTNNKATVKFDI 269
QY 129 TQQAASLLELAQKVGIEIV--YPDVASSSSFTVFGIYQQLQSSSAVDIKASDLPQ 185
D 270 TNNAEQLLNQANI--MQLNTQCPYVSTNNENTPGGGQPGWGLSTSGNAGSI----FQQ 323
QY 186 SGDOVNKDITQKRYRTILDSTVVASQREYINSVKQKQPISNYY-VGYSRMEIKDIIRDQ 244
D 324 EFSQVTSMTKKAQELIAQSKIVSENAQNNNDTGKPPNPPTDASFAQSM-----LKNA 377
QY 245 QYNVOLIGTSKPYVYTDVLAALNSNLCDK-----KQYAVEVIXKL----- 286
D 378 QAQEMFNLSPQVKKNLSEVMNNNNVNEKLAGFGKEEYMTNFVSAFLASCKDGGTLPVAG 437
QY 287 --TTLVLVDLIGLGLTLPANKNGIAHLAKSSNFYAQLSQQPDAKSESEVRLCVDPANKE 343
D 438 VTSNTWGACAYVGETISALTNISIAH-----FGTDEQIQQAENIADTL--VNFKSR 487
```

```

RESULT 2
US-09-951-402-3
; Sequence 3, Application US/09951402
; Patent No. US20020168752A1
; GENERAL INFORMATION:
; APPLICANT: Bartel, Paul L.
; TITLE OF INVENTION: MMSC2- An MMAC1 Interacting Protein
; FILE REFERENCE: MMSC2
; CURRENT APPLICATION NUMBER: US/09/951,402
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 09/306,998
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 60/084,740
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2037
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-951-402-3
```

```

Query Match          5.4%; Score 97.5; DB 9; Length 2037;
Best Local Similarity 21.4%; Pred. No. 11;
Matches 75; Conservative 48; Mismatches 147; Indels 81; Gaps 15;
```

```

QY 7 TLTVGLFPYLPMSNNGNEVKLINLIKDVLPYQVSGYNIETEPDQSDASLQSLPQVFS 66
D 778 TVRIGVAKPLPLSPREG---YVSAKED-----SFLYPPHSC--EAGLADKP-LFR 822
QY 67 TDSIFLPLVLSLGVKSLDESIVRGVTDLHSFVSSASVNGSVYGPQYLCNFFLLSP 126
D 823 ADLA---LVGTNDADLVDES-----TFSPSPSPENDSISTQASLISHGSSCG 868
QY 127 NGTQOASSLLELAQKVGIEQIVYPDVASSSSFTVFGIY-----QQLQSSSSAAVDI 178
D 869 DGLNVGSSLPSPPK-----DVLENSCDPVLDMHSLDEELYTONLQRODENTPSV 919
```

```

QY 179 KASDLPQSGQOVN-----KQITQKRYRTILDSTVVASQREYINSVKQKQPISNYYVGYSSES 233
D 920 DISMGPASGFTINDYTPANAIEQYE--CENTIWTESHLPSEV-----IS 963
QY 234 MCEIKDIIIPQOQVNVQILGTSKPYVYTDVLAALNSNLCDKQKQYAVEVIXKLTLNTVLVD 293
D 964 SAEIPLSVLPDS-----AGKSEYILQSSSLACNAE-CWMLQNVSKSPFER--TINIAGK 1014
QY 294 LLAGLGLTLPANKNGIAHLAKSSNFYAQLSQQPDAKSESEVRLCVDPANKE 344
D 1015 NSSLGMVTSANRKGILGMIVRSIIHGAIIS-----RDRIRIAGDCLISINEE 1060
```

```

RESULT 3
US-09-951-401-3
; Sequence 3, Application US/09951401
; Patent No. US2002015104A1
; GENERAL INFORMATION:
; APPLICANT: Bartel, Paul L.
; APPLICANT: Tavtigian, Sean V.
; TITLE OF INVENTION: MMSC2- An MMAC1 Interacting Protein
; FILE REFERENCE: MMSC2
; CURRENT APPLICATION NUMBER: US/09/951,401
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 09/306,998
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 60/084,740
; PRIOR FILING DATE: 1999-05-08
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2037
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-951-401-3
```

```

Query Match          5.4%; Score 97.5; DB 10; Length 2037;
Best Local Similarity 21.4%; Pred. No. 11;
Matches 75; Conservative 48; Mismatches 147; Indels 81; Gaps 15;
```

```

QY 7 TLTVGLFPYLPMSNNGNEVKLINLIKDVLPYQVSGYNIETEPDQSDASLQSLPQVFS 66
D 778 TVRIGVAKPLPLSPREG---YVSAKED-----SFLYPPHSC--EAGLADKP-LFR 822
QY 67 TDSIFLPLVLSLGVKSLDESIVRGVTDLHSFVSSASVNGSVYGPQYLCNFFLLSP 126
D 823 ADLA---LVGTNDADLVDES-----TFSPSPSPENDSISTQASLISHGSSCG 868
QY 127 NGTQOASSLLELAQKVGIEQIVYPDVASSSSFTVFGIY-----QQLQSSSSAAVDI 178
D 869 DGLNVGSSLPSPPK-----DVLENSCDPVLDMHSLDEELYTONLQRODENTPSV 919
QY 179 KASDLPQSGQOVN-----KQITQKRYRTILDSTVVASQREYINSVKQKQPISNYYVGYSSES 233
D 920 DISMGPASGFTINDYTPANAIEQYE--CENTIWTESHLPSEV-----IS 963
QY 234 MCEIKDIIIPQOQVNVQILGTSKPYVYTDVLAALNSNLCDKQKQYAVEVIXKLTLNTVLVD 293
D 964 SAEIPLSVLPDS-----AGKSEYILQSSSLACNAE-CWMLQNVSKSPFER--TINIAGK 1014
QY 294 LLAGLGLTLPANKNGIAHLAKSSNFYAQLSQQPDAKSESEVRLCVDPANKE 344
D 1015 NSSLGMVTSANRKGILGMIVRSIIHGAIIS-----RDRIRIAGDCLISINEE 1060
```

```

RESULT 4
US-09-922-101-3
; Sequence 3, Application US/09922101
; Patent No. US20020146711A1
; GENERAL INFORMATION:
; APPLICANT: Bartel, Paul L.
; APPLICANT: Tavtigian, Sean V.
; TITLE OF INVENTION: MMSC2- An MMAC1 Interacting Protein
```

FILE REFERENCE: MMSC2  
CURRENT APPLICATION NUMBER: US/09/922.101  
CURRENT FILING DATE: 2001-08-06  
PRIOR APPLICATION NUMBER: 09/306,998  
PRIOR FILING DATE: 1999-05-07  
NUMBER OF SEQ ID NOS: 72  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 2037  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-922-101-3

Query Match 5.4%; Score 97.5; DB 10; Length 2037;  
Best Local Similarity 21.4%; Pred. No. 11;  
Matches 75; Conservative 48; Mismatches 147; Indels 81; Gaps 15;

QY 7 TLTVGLPEPLPSMNGENGVKLNLKQVLPQVSGVNIETEPDQSDASGLPDVFS 66  
D 778 TVRIGVAKPLPESPEG-----YVSAKED-----SFLYPHSC--EEAGLADKP-LFR 822  
QY 67 TDSIFLPYLVSGVKSLSDELVGVGTGDLHSPVSSASVNGSVYGFPOYLCSNPLSSP 126  
D 823 ADLA-----LVGTNDADLVDS-----TFSPYSPENDSITSTQASITLSHGSSCG 868  
QY 127 NGTQOASSLLELAQKVGVEQIVPDAVSSSSFTVFGLY-----QQLQSSSSAAVDI 178  
D 869 DGLNVGSSLPSPPK-----DVIENSCDPLDLHMSLELYTQNLOROBENTPSV 919  
QY 179 KASLPQSGDQV-----KDTQKRTLLDSTVVASQREVINSVKQKPISNVYVGSSES 233  
D 920 DISMGPAISGFINDYTPANAIEQQYE--CENTIIVTESHLSSEV-----IS 963  
QY 234 MCEIKDIIRDQVNVQLIGTSDKPYVYTDVLAHNSNLCDEKQXAVEYIKXLLTNTLVLD 293  
D 964 SAELPSVLPDS-----AGKSEYLLBOSSLACNAE--CVMQNVSKSEFER--TINIAGK 1014  
QY 294 LLAGLGLTLPANKNGIAHLAKSSNFYAOISQOPDAKESFVRVLRCDVPAKKE 344  
D 1015 NSSLGMTVASANKDGLGMIIVRSIIHGNAIS-----RDGRIALGDCILSINEE 1060

RESULT 5  
US-09-922-598-194  
Sequence 194, Application US/09992598  
Patent No. US20020160384A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnovers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerltsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Grimaldi, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kijavini, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Matanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
TITLE OF INVENTION: Acids Encoding the Same

FILE REFERENCE: P2730PIC20  
CURRENT APPLICATION NUMBER: US/09/992,598  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088555  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11

[illegible]

```

      PRIOR FILING DATE: 1998-06-25
      PRIOR APPLICATION NUMBER: 60/090696
      PRIOR FILING DATE: 1998-06-25
      PRIOR APPLICATION NUMBER: 60/090862
      PRIOR FILING DATE: 1998-06-26
      PRIOR APPLICATION NUMBER: 60/090863
      PRIOR FILING DATE: 1998-06-26
      PRIOR APPLICATION NUMBER: 60/091360
      PRIOR FILING DATE: 1998-07-01
      PRIOR APPLICATION NUMBER: 60/091478
      PRIOR FILING DATE: 1998-07-02
      PRIOR APPLICATION NUMBER: 60/091544
      PRIOR FILING DATE: 1998-07-01
      PRIOR APPLICATION NUMBER: 60/091519
      PRIOR FILING DATE: 1998-07-02
      PRIOR APPLICATION NUMBER: 60/091626
      PRIOR FILING DATE: 1998-07-02
      PRIOR APPLICATION NUMBER: 60/091633
      PRIOR FILING DATE: 1998-07-02
      PRIOR APPLICATION NUMBER: 60/091978
      PRIOR FILING DATE: 1998-07-07
      PRIOR APPLICATION NUMBER: 60/091982
      PRIOR FILING DATE: 1998-07-07
      PRIOR APPLICATION NUMBER: 60/092182
      PRIOR FILING DATE: 1998-07-09

Query Match          5.4%; Score 97; DB 9; Length 615;
Best Local Similarity 21.3%; Pred. No. 2.3;
Matches   54; Conservative    50; Mismatches   112; Indels   38; Gaps   10;

QY  119 SNFLSSPNGTQAASSLLELAOKVGEQIYYPDVAASSSFVFGLYOQLDSSSSAAVD- 177
DB  147 AGIIIVSVNPAQOMELIEYVKVGKCALVFPKOFKTOQY--YNVLKOICEVEVIAORGCA 204
QY  178 IKASDLPOSGDOVNKDITOKRYRILIDSTVVA-SOREYINSYCKPKPISNYVYGSENC 236
DB  205 LKSGRLPEDLTTVISVDAPRLGCTLLDDEVNAAGSTRCHLDQO-----YNQGFLS 253
QY  237 IKDIIRDQYNNOLI-GTSDKPRVYT---DYALNSMLCEKOKV--AVEIKNLTLNT 289
DB  254 CHDP-----NIPTSGTGSPKCATLSHVINVNNSNLTGERLKLHEKTPEQLRMILBNP 308
QY  290 LVEDDLGLGLTPANKNGIAHLAKSSNFYAOLSOQFPAKEE-----VRLRCVD 339
DB  309 LYHCIGSVAQTMCMLMGATLLIASPIFNCRKALEAISREGTFLYXTPTMFVIDLNOPD 368
QY  340 FANKEVKN-CAGVL 352
DB  369 FSYTDISTMCGGYI 382

RESULT 6
US-09-989--293A-194
Sequence 194, Application US/09989293A
Patent No. US20020177164A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gottard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gueney, Austin L.
APPLICANT: Kijavini, Ivar A.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
```

APPLICANT: Roy, Margaret Ann  
 APPLICANT: Stewart, Timothy A.  
 APPLICANT: Tumas, Daniel  
 APPLICANT: Watanabe, Colin K.  
 APPLICANT: Williams, P. Mickey  
 APPLICANT: Wood, William I.  
 APPLICANT: Zheng, Zemin  
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 FILE REFERENCE: P2730PIC66  
 CURRENT APPLICATION NUMBER: US/09/989,293A  
 PRIOR FILING DATE: 2001-11-20  
 PRIOR APPLICATION NUMBER: 60/049787  
 PRIOR FILING DATE: 1997-06-16  
 PRIOR APPLICATION NUMBER: 60/062250  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/065186  
 PRIOR FILING DATE: 1997-11-12  
 PRIOR APPLICATION NUMBER: 60/065311  
 PRIOR FILING DATE: 1997-11-13  
 PRIOR APPLICATION NUMBER: 60/066770  
 PRIOR FILING DATE: 1997-11-24  
 PRIOR APPLICATION NUMBER: 60/075945  
 PRIOR FILING DATE: 1998-02-25  
 PRIOR APPLICATION NUMBER: 60/078910  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/083322  
 PRIOR FILING DATE: 1998-04-28  
 PRIOR APPLICATION NUMBER: 60/084600  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/087106  
 PRIOR FILING DATE: 1998-05-28  
 PRIOR APPLICATION NUMBER: 60/087607  
 PRIOR FILING DATE: 1998-06-02  
 PRIOR APPLICATION NUMBER: 60/087609  
 PRIOR FILING DATE: 1998-06-02  
 PRIOR APPLICATION NUMBER: 60/087759  
 PRIOR FILING DATE: 1998-06-02  
 PRIOR APPLICATION NUMBER: 60/087827  
 PRIOR FILING DATE: 1998-06-03  
 PRIOR APPLICATION NUMBER: 60/088021  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088025  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088026  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088028  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088029  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088030  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088033  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088326  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088167  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088202  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088212  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088217  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088655  
 PRIOR FILING DATE: 1998-06-09  
 PRIOR APPLICATION NUMBER: 60/088734  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088738  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088742  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088810

PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088824  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088826  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088858  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/088861  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/088876  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/089105  
 PRIOR FILING DATE: 1998-06-12  
 PRIOR APPLICATION NUMBER: 60/089440  
 PRIOR FILING DATE: 1998-06-16  
 PRIOR APPLICATION NUMBER: 60/089512  
 PRIOR FILING DATE: 1998-06-16  
 PRIOR APPLICATION NUMBER: 60/089514  
 PRIOR FILING DATE: 1998-06-16  
 PRIOR APPLICATION NUMBER: 60/089532  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089538  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089598  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089599  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089600  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089653  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089801  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089907  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089908  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089947  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/089948  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/089952  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/090246  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090252  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090254  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090349  
 PRIOR FILING DATE: 1998-06-23  
 PRIOR APPLICATION NUMBER: 60/090355  
 PRIOR FILING DATE: 1998-06-23  
 PRIOR APPLICATION NUMBER: 60/090429  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090431  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090435  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090444  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090445  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090472  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090535  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090540  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090542  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090557  
 PRIOR FILING DATE: 1998-06-24

PRIOR APPLICATION NUMBER: 60/090676  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090678  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090690  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090694  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090695  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090696  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090862  
 PRIOR FILING DATE: 1998-06-26  
 PRIOR APPLICATION NUMBER: 60/090863  
 PRIOR FILING DATE: 1998-06-26  
 PRIOR APPLICATION NUMBER: 60/091360  
 PRIOR FILING DATE: 1998-07-01  
 PRIOR APPLICATION NUMBER: 60/091478  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091544  
 PRIOR FILING DATE: 1998-07-01  
 PRIOR APPLICATION NUMBER: 60/091519  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091626  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091633  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091978  
 PRIOR FILING DATE: 1998-07-07  
 PRIOR APPLICATION NUMBER: 60/091982  
 PRIOR FILING DATE: 1998-07-07  
 PRIOR APPLICATION NUMBER: 60/092182  
 PRIOR FILING DATE: 1998-07-09

Query Match 5.4%, Score 97, DB 9, length 615;  
 Best Local Similarity 21.3%, Pred. No. 2.3, Matches 54, Conservative 50, Mismatches 112, Indels 38, Gaps 10;

QY 119 SNFLSPNGTQOASSLELAQKVGEQIVYDVASSSSFTYFGLYOOLQSSSAVD- 177  
 Db 147 AGILIVSNPAPQAMELEVLYLKVCKALVPPKQKTOQY--YVNLKQICPEVENAOGCA 204  
 QY 178 IKASDLPOSGQOVNDITQKRTTILDSVVA--SQREYINSVKQGPISNYYVGSSENC 236  
 Db 205 LKSGRLPLTLTYISVDADPLGLLDEVVAAGSTRQHLDQLQ-----YNOQLS 253  
 QY 237 IKDIIRDQVNVQL-GTSDKRYVT---DVIALNSNLCDEKQY---AVEYIKULLNT 289  
 Db 254 CHDPT-----NIQFTSGTTSKRGATLSHYNIUNNSNLTGERTKLHEKTPQQLRMILPNP 308  
 QY 290 IVLDLGLGLTPANKNGIAHLAKSSNFYALSOQOFDAKESB-----VRVLRCVD 339  
 Db 309 LYHCGSVAGTMCMLMYGATLILASPIFGKALALASRERGTFLYGPPTMVDILNPD 368  
 QY 340 FANKEVKN-CAGVL 352  
 Db 369 FSSYDSTMCGGVI 382

RESULT 7  
 US-09-989-735-194  
 Sequence 194, Application US/09989735  
 Publication No. US20020193299A1  
 GENERAL INFORMATION:  
 APPLICANT: Ashkenazi, Avi J.  
 APPLICANT: Baker, Kevin P.  
 APPLICANT: Botstein, David  
 APPLICANT: Deenoyers, Luc  
 APPLICANT: Batson, Dan L.  
 APPLICANT: Ferrara, Napoleone  
 APPLICANT: Fong, Sherman  
 APPLICANT: Gerber, Hanspeter

APPLICANT: Gerltsen, Mary E.  
 APPLICANT: Goddard, Audrey  
 APPLICANT: Godowski, Paul J.  
 APPLICANT: Grimaldi, J. Christopher  
 APPLICANT: Gurney, Austin L.  
 APPLICANT: Kijavini, Ivar J.  
 APPLICANT: Napier, Mary A.  
 APPLICANT: Pan, James  
 APPLICANT: Paoni, Nicholas F.  
 APPLICANT: Roy, Margaret Ann  
 APPLICANT: Stewart, Timothy A.  
 APPLICANT: Tumas, Daniel  
 APPLICANT: Watanabe, Colin K.  
 APPLICANT: Williams, P. Mickey  
 APPLICANT: Wood, William I.  
 APPLICANT: Zhang, Zemin  
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 TITLE OF INVENTION: Acids Encoding the Same  
 FILE REFERENCE: P2730P1C61  
 CURRENT APPLICATION NUMBER: US/09/989,735  
 CURRENT FILING DATE: 2001-11-19  
 PRIOR APPLICATION NUMBER: 60/049787  
 PRIOR FILING DATE: 1997-06-16  
 PRIOR APPLICATION NUMBER: 60/062250  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/065186  
 PRIOR FILING DATE: 1997-11-12  
 PRIOR APPLICATION NUMBER: 60/065311  
 PRIOR FILING DATE: 1997-11-13  
 PRIOR APPLICATION NUMBER: 60/066770  
 PRIOR FILING DATE: 1997-11-24  
 PRIOR APPLICATION NUMBER: 60/075945  
 PRIOR FILING DATE: 1998-02-25  
 PRIOR APPLICATION NUMBER: 60/078910  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/083322  
 PRIOR FILING DATE: 1998-04-28  
 PRIOR APPLICATION NUMBER: 60/084600  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/087106  
 PRIOR FILING DATE: 1998-05-28  
 PRIOR APPLICATION NUMBER: 60/087607  
 PRIOR FILING DATE: 1998-06-02  
 PRIOR APPLICATION NUMBER: 60/087609  
 PRIOR FILING DATE: 1998-06-02  
 PRIOR APPLICATION NUMBER: 60/087759  
 PRIOR FILING DATE: 1998-06-02  
 PRIOR APPLICATION NUMBER: 60/087827  
 PRIOR FILING DATE: 1998-06-03  
 PRIOR APPLICATION NUMBER: 60/088021  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088025  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088026  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088028  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088029  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088030  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088033  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088326  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088167  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088202  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088212  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088217  
 PRIOR FILING DATE: 1998-06-05

PRIOR APPLICATION NUMBER: 60/088655	
PRIOR FILING DATE: 1998-06-09	
PRIOR APPLICATION NUMBER: 60/088733	
PRIOR FILING DATE: 1998-06-10	
PRIOR APPLICATION NUMBER: 60/088738	
PRIOR FILING DATE: 1998-06-10	
PRIOR APPLICATION NUMBER: 60/088742	
PRIOR FILING DATE: 1998-06-10	
PRIOR APPLICATION NUMBER: 60/088810	
PRIOR FILING DATE: 1998-06-10	
PRIOR APPLICATION NUMBER: 60/088822	
PRIOR FILING DATE: 1998-06-10	
PRIOR APPLICATION NUMBER: 60/088823	
PRIOR FILING DATE: 1998-06-10	
PRIOR APPLICATION NUMBER: 60/088856	
PRIOR FILING DATE: 1998-06-11	
PRIOR APPLICATION NUMBER: 60/088861	
PRIOR FILING DATE: 1998-06-11	
PRIOR APPLICATION NUMBER: 60/088876	
PRIOR FILING DATE: 1998-06-11	
PRIOR APPLICATION NUMBER: 60/089105	
PRIOR FILING DATE: 1998-06-12	
PRIOR APPLICATION NUMBER: 60/089444	
PRIOR FILING DATE: 1998-06-16	
PRIOR APPLICATION NUMBER: 60/089512	
PRIOR FILING DATE: 1998-06-16	
PRIOR APPLICATION NUMBER: 60/089514	
PRIOR FILING DATE: 1998-06-16	
PRIOR APPLICATION NUMBER: 60/089523	
PRIOR FILING DATE: 1998-06-17	
PRIOR APPLICATION NUMBER: 60/089538	
PRIOR FILING DATE: 1998-06-17	
PRIOR APPLICATION NUMBER: 60/089598	
PRIOR FILING DATE: 1998-06-17	
PRIOR APPLICATION NUMBER: 60/089599	
PRIOR FILING DATE: 1998-06-17	
PRIOR APPLICATION NUMBER: 60/089600	
PRIOR FILING DATE: 1998-06-17	
PRIOR APPLICATION NUMBER: 60/089653	
PRIOR FILING DATE: 1998-06-17	
PRIOR APPLICATION NUMBER: 60/089801	
PRIOR FILING DATE: 1998-06-18	
PRIOR APPLICATION NUMBER: 60/089907	
PRIOR FILING DATE: 1998-06-18	
PRIOR APPLICATION NUMBER: 60/089908	
PRIOR FILING DATE: 1998-06-18	
PRIOR APPLICATION NUMBER: 60/089944	
PRIOR FILING DATE: 1998-06-19	
PRIOR APPLICATION NUMBER: 60/089948	
PRIOR FILING DATE: 1998-06-19	
PRIOR APPLICATION NUMBER: 60/089952	
PRIOR FILING DATE: 1998-06-19	
PRIOR APPLICATION NUMBER: 60/090025	
PRIOR FILING DATE: 1998-06-22	
PRIOR APPLICATION NUMBER: 60/090254	
PRIOR FILING DATE: 1998-06-22	
PRIOR APPLICATION NUMBER: 60/090349	
PRIOR FILING DATE: 1998-06-23	
PRIOR APPLICATION NUMBER: 60/090355	
PRIOR FILING DATE: 1998-06-23	
PRIOR APPLICATION NUMBER: 60/090429	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090431	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090435	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090445	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090455	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090472	
PRIOR FILING DATE: 1998-06-24	

```

      PRIOR FILING DATE: 1998-06-24
      PRIOR APPLICATION NUMBER: 60/090535
      PRIOR FILING DATE: 1998-06-24
      PRIOR APPLICATION NUMBER: 60/090540
      PRIOR FILING DATE: 1998-06-24
      PRIOR APPLICATION NUMBER: 60/090542
      PRIOR FILING DATE: 1998-06-24
      PRIOR APPLICATION NUMBER: 60/090557
      PRIOR FILING DATE: 1998-06-24
      PRIOR APPLICATION NUMBER: 60/090676
      PRIOR FILING DATE: 1998-06-25
      PRIOR APPLICATION NUMBER: 60/090678
      PRIOR FILING DATE: 1998-06-25
      PRIOR APPLICATION NUMBER: 60/090690
      PRIOR FILING DATE: 1998-06-25
      PRIOR APPLICATION NUMBER: 60/090694
      PRIOR FILING DATE: 1998-06-25
      PRIOR APPLICATION NUMBER: 60/090695
      PRIOR FILING DATE: 1998-06-25
      PRIOR APPLICATION NUMBER: 60/090696
      PRIOR FILING DATE: 1998-06-25
      PRIOR APPLICATION NUMBER: 60/090862
      PRIOR FILING DATE: 1998-06-26
      PRIOR APPLICATION NUMBER: 60/090863
      PRIOR FILING DATE: 1998-06-26
      PRIOR APPLICATION NUMBER: 60/091360
      PRIOR FILING DATE: 1998-07-01
      PRIOR APPLICATION NUMBER: 60/091478
      PRIOR FILING DATE: 1998-07-02
      PRIOR APPLICATION NUMBER: 60/091544
      PRIOR FILING DATE: 1998-07-01
      PRIOR APPLICATION NUMBER: 60/091519
      PRIOR FILING DATE: 1998-07-02
      PRIOR APPLICATION NUMBER: 60/091626
      PRIOR FILING DATE: 1998-07-02
      PRIOR APPLICATION NUMBER: 60/091633
      PRIOR FILING DATE: 1998-07-02
      PRIOR APPLICATION NUMBER: 60/091978
      PRIOR FILING DATE: 1998-07-07
      PRIOR APPLICATION NUMBER: 60/091982
      PRIOR FILING DATE: 1998-07-07
      PRIOR APPLICATION NUMBER: 60/092182
      PRIOR FILING DATE: 1998-07-09

Query Match      5.4%; Score 97; DB 9; Length 615;
Best Local Similarity 21.3%; Pred. No. 2.3;
Matches 54; Conservative 50; Mismatches 112; Indels 38; Gaps 10;

OY 119 SNFLSSPENGTOQASLSLELAKQYGYEDIYVPDYVASSSPFVFGLYQOULLOSSSSAAND- 177
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 147 AGILVSNPAPYQAMELELYVLKKGVCCKALVPKQFQTOY--YNNLKOICEVEVAOGCA 204
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 178 IKASDLPQSSGOQVKKDITQKXRTILDSIVVA--SQREYINSVKQCKPISINYVGSSEMC 236
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 205 LKSGRPLPDLTVIVSDAPLPETLLDDEVVAAGSTRQHDLOLO-----YNOQFLS 253
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 237 IKDIIIPDOQVNWQLI--GTSDKRPYYVT---DYVALNSNLCDKQKV---AVEYIKNLLTNT 289
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 254 CHDPI-----NIQTSGTGTSFKGATLSHNYIVNNSNLIGERLKHKEKTPEDGLRWILNPN 308
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 290 LVLDLGLGLTLTPANKNGIAHLAKSSNFYAOLSOQFPAKES-----VRLRCVD 339
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 309 LYHCGISVAGTMCMCLMGATLILASPIFNKGKALEAISREGRTPFYGTPTMEVDILNOPD 368
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 340 FANKEVKN-CAGVL 352
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 369 FSSYDISTMGCVI 382
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 8
US-09-990-444-194
Sequence 194, Application US/09990444
Publication No. US20020193100A1

```

GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gertsens, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kijavini, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tamas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zheng, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730P1C19  
CURRENT APPLICATION NUMBER: US/09/990,444  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326

PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24

PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090535  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090540  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090542  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090676  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090678  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090690  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090694  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090695  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090696  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090862  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091478  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091544  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091626  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091633  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091978  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/092182  
PRIOR FILING DATE: 1998-07-09

Query Match 5.4%; Score 97; DB 9; Length 615;  
Best Local Similarity 21.3%; Pred. No. 2.3;  
Matches 54; Conservative 50; Mismatches 112; Indels 38; Gaps 10;

QY 119 SNFLSSPNGTQOASLLLEAOKYGEIOIYPPDVASSSPFVGLYQOULLOSSSAVD- 177  
DB 147 AGILIVNPAVYQAMELEYVLKKVGCALVFPKQFKTOY--YVWLKQICEVENAPGCA 204  
QY 178 IKASDLPOSGDQVKNKDTQKRYTLDSTVVA--SOREYVNSYKQKPISNVYVGSSEACE 236  
DB 205 LKSGRLPDLTIVVDAPLPGTLLDEYVAGSTRQHLDOUQ-----YNOQPLS 253  
QY 237 IKDIIRDQYVQLI-GTSDKPYVYT--DVLALNSMLCDEKQV--AVEVINKLLTNT 289  
DB 254 CHDEI-----NIQFTSGTSGPKGATLSHYNIYVNSNLTGSRLLKLTHEKTPQRLMILPNP 308  
QY 290 LVLDLGLGLTPANKGIAIALAKSSNFYALSOQFDKKEE-----VRLKQVD 339  
DB 309 LYHCLGSSVAGTMCIMYATLILASPIFNKKALEAISREGRGFLYGTPTMVFIDLNPDP 368

QY 340 FANKEVKN-CAGVL 352  
DB 369 FSSYDISTMCGVIL 382

RESULT 9  
US-09-989-730-194  
Sequence 194, Application US/09989730  
Publication No. US20020197674A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Bolstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kijavlin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2730PLC69  
CURRENT APPLICATION NUMBER: US/09/989,730  
CURRENT FILING DATE: 2001-11-20  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04



QY 237 IKDIIRDQOYVWOLI-GTSDKPYVYT---DYALNSNLCDKOKY---AVEVKNLLTNT 289  
Db 254 CHDPI-----NIQFTSGTGTSPKCATLISHYNI VVNSNLSIGRLKHEKTPQGLMILPNP 308  
QY 290 LVLDLGLGLTPANKKGIHIAKSSNFYQOLSQOFDAKES-----YAVLRQVD 339  
Db 309 LYHCLGSVAGTMCMLWATLILASPIFNCKKALEAISREGRFTLYGTPTMFVILNPD 368  
QY 340 FANKEVKN-CAGVL 352  
Db 369 FSSYDISTMCGVI 382

RESULT 10  
US-09-990-436-194  
; Sequence 194, Application US/09990436  
; Publication No. US20020198148A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gertlisen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kijavlin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2730P1C14  
; CURRENT APPLICATION NUMBER: US/09/990,436  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759

; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088021  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088030  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088858  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088861  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088876  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/089105  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089440  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089512  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089514  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089532  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089598  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089599  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089600  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089653  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089907  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089908  
; PRIOR FILING DATE: 1998-06-18





;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

Query Match 5.4%; Score 97; DB 9; Length 615;  
Best Local Similarity 21.3%; Pred. No. 2.3;  
Matches 54; Conservative 50; Mismatches 112; Indels 38; Gaps 10;

Qy 119 SNFLSSPNGQQAASLLEAQKGYEDIVYVDVAASSSFVFGLYQQLDSSSSAAND- 177  
Db 147 AGILTVSNPAPYQAMELELYKVKCKALVFPKQFKTOY--YVNLKQICEVEVAQGA 204  
Qy 178 IKASDLPOSGDQVNDITQKRTILDSTVVA--SOREYINSVKQKGFISNYVYGSSEMC 236  
Db 205 LKSGRLDPLTVIVSDAPLPGTLLDDEVVAAGSTRQHLDQO-----YNQQLS 253  
Db 237 IKDIIRDQVNVQLI-GTSDKPYVYT--DVLALNSLCKDEKQV--AVEVINKLTLNT 289  
Db 254 CHDPI-----NIQFTSGTGSFKGATLSHYNI VNNNSNITLGRLLKHEKTPQQLRMILPNP 308  
Qy 290 LVLDLLGLGLTPANKGIAHLAKSSNRYAQLSQQFDKKESE-----VRYLRQVD 339  
Db 309 LYHGLGVAGTMMCLMGATLLASPIFNKKALEAISREKGTFLYGTPTMFVILNPD 368  
Qy 340 FANKEVKN-CAGVL 352  
Db 369 FSSYDIDTMCGSVI 382

RESULT 12  
US-09-993-687-194  
; Sequence 194, Application US/09993687  
; Publication No. US20020198149A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin J.  
; APPLICANT: Kijavini, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2730PIC11  
; CURRENT APPLICATION NUMBER: US/09/993.687  
; CURRENT FILING DATE: 2002-11-14  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311

;; PRIOR FILING DATE: 1997-11-13  
;; PRIOR APPLICATION NUMBER: 60/066770  
;; PRIOR FILING DATE: 1997-11-24  
;; PRIOR APPLICATION NUMBER: 60/075945  
;; PRIOR FILING DATE: 1998-02-25  
;; PRIOR APPLICATION NUMBER: 60/078910  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/083322  
;; PRIOR FILING DATE: 1998-04-28  
;; PRIOR APPLICATION NUMBER: 60/084600  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/087106  
;; PRIOR FILING DATE: 1998-05-28  
;; PRIOR APPLICATION NUMBER: 60/087607  
;; PRIOR FILING DATE: 1998-06-02  
;; PRIOR APPLICATION NUMBER: 60/087609  
;; PRIOR FILING DATE: 1998-06-02  
;; PRIOR APPLICATION NUMBER: 60/087759  
;; PRIOR FILING DATE: 1998-06-02  
;; PRIOR APPLICATION NUMBER: 60/087827  
;; PRIOR FILING DATE: 1998-06-03  
;; PRIOR APPLICATION NUMBER: 60/088021  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088025  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088026  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088028  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088029  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088030  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088033  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088326  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088167  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088202  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088212  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088217  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088555  
;; PRIOR FILING DATE: 1998-06-09  
;; PRIOR APPLICATION NUMBER: 60/088734  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088738  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088742  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088810  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088824  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088826  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088858  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/088861  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/088876  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/089105  
;; PRIOR FILING DATE: 1998-06-12  
;; PRIOR APPLICATION NUMBER: 60/089440  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089512  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089514  
;; PRIOR FILING DATE: 1998-06-16

PRIOR APPLICATION NUMBER: 60/089532  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089538  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089598  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089599  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089600  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089653  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089801  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089907  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089908  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089947  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/089948  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/089952  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/090246  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090252  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090254  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090349  
 PRIOR FILING DATE: 1998-06-23  
 PRIOR APPLICATION NUMBER: 60/090355  
 PRIOR FILING DATE: 1998-06-23  
 PRIOR APPLICATION NUMBER: 60/090429  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090431  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090435  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090444  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090445  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090472  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090535  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090540  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090542  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090557  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090676  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090678  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090690  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090694  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090695  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090696  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090862  
 PRIOR FILING DATE: 1998-06-26  
 PRIOR APPLICATION NUMBER: 60/090863  
 PRIOR FILING DATE: 1998-06-26  
 PRIOR APPLICATION NUMBER: 60/091360  
 PRIOR FILING DATE: 1998-07-01  
 PRIOR APPLICATION NUMBER: 60/091478

PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091544  
 PRIOR FILING DATE: 1998-07-01  
 PRIOR APPLICATION NUMBER: 60/091519  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091626  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091633  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091978  
 PRIOR FILING DATE: 1998-07-07  
 PRIOR APPLICATION NUMBER: 60/091982  
 PRIOR FILING DATE: 1998-07-07  
 PRIOR APPLICATION NUMBER: 60/092182  
 PRIOR FILING DATE: 1998-07-09

Query Match 5.4%; Score 97; DB 9; length 615;  
 Best Local Similarity 21.3%; Pred. No. 2.3; Indels 38; Gaps 10;  
 Matches 54; Conservative 50; Mismatches 112;

QY 119 SNFLSSPNGTQOASSLLELAQKGYEQIYVPDVASSSFTVEGLYQOQLQSSSSAAVD- 177  
 DB 147 AGITLVSNPAPYQAMELEYUKVGCALVPPKQFKQY-YNVLKQICEVENAQPQA 204  
 QY 178 IKASDLPOSGDQVKNKDTQKXRTITDSTVVA-SQREYINSVKQKPISNYYVGSSEMC 236  
 DB 205 LKSGRLPDLTTVTSVDAPLPGTLLDEVAAGSTRHLDLQ-----YNGQFLS 253  
 QY 237 IKDIIRDQYNNVOLI-GTSDEPYVT--DYLAISNLCDEKQV--ANEVINLLTNT 289  
 DB 254 CHDPI-----NIQTSQTSSPKGATLSHYNIWNNSNIIERLTHKRTBOELMIIPNP 308  
 QY 290 LVLDLGLGLTPANKNGIAHIAKSSNFYAQLSQQFPAKSE-----VRLACVD 339  
 DB 309 LYHCLGSVAGTMCMLGATLILASPFGKALDAISRGRGFLYGTPTMFWIDLNP 368  
 QY 340 FANKEVKN-CAGVL 352  
 DB 369 FSSYDISTMCGVI 382

RESULT 13  
 US-09-989-734-194  
 Sequence 194, Application US/09989734  
 Publication No. US20030003531A1  
 GENERAL INFORMATION:  
 APPLICANT: Ashkenazi, Avi J.  
 APPLICANT: Baker, Kevin P.  
 APPLICANT: Botstein, David  
 APPLICANT: Desnoyers, Luc  
 APPLICANT: Baton, Dan L.  
 APPLICANT: Ferrara, Napoleone  
 APPLICANT: Fong, Sherman  
 APPLICANT: Gerber, Hanspeter  
 APPLICANT: Gerlitsen, Mary E.  
 APPLICANT: Goddard, Audrey  
 APPLICANT: Godowski, Paul J.  
 APPLICANT: Grimaldi, J. Christopher  
 APPLICANT: Gurney, Austin L.  
 APPLICANT: Kljavin, Ivar J.  
 APPLICANT: Nadier, Mary A.  
 APPLICANT: Pan, James  
 APPLICANT: Paoni, Nicholas F.  
 APPLICANT: Roy, Margaret Ann  
 APPLICANT: Stewart, Timothy A.  
 APPLICANT: Tumas, Daniel  
 APPLICANT: Watanabe, Colin K.  
 APPLICANT: Williams, P. Mickey  
 APPLICANT: Wood, William I.  
 APPLICANT: Zhang, Zemin  
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 TITLE OF INVENTION: Acids Encoding the Same  
 FILE REFERENCE: P2730P1C64

[illegible]

```

; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

```

```

Query Match          5.4%; Score 97; DB 9; Length 615;
Best Local Similarity 21.3%; Pred. No. 2.3;
Matches 54; Conservative 50; Mismatches 112; Indels 38; Gaps 10;

```

```

QY 119 SNFLSSPNTQQAASLLELAQKGYEQIYVPDVAASSTFTVFGLYQQLLOSSSAVD- 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 147 AGITVSNPAVQAAMELEVVLKKVGCCKALVFPKPKQY--YVNLKQICEVENAOGCA 204
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 178 IKASDLQSGQVVKDITQKRTILDSTVVA--SQREYINSKQGGKPISNYVGSSESC 236
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 205 LKSQLPLDITVISDAPLPGTLLDEVVAGSTRQHLDQ-----YNQOPLS 253
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 237 IKDIIRDOQVNVQLI-GTSDKPYVYT--DVALNSNLDKQKV--AVEVKNLTLNT 289
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 254 CHDPI-----NIQFTSGTSSPKGATLSHYVIVNNSNLGERLKLHKTPEQKMLPNP 308
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 290 LVLDLGLGLTLTPANKNGIAHLAKSNFYAQLSQOFPDASE-----VRYLRCDV 339
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 309 LYNHCLGSVAGTMMCLMGCATLILASPIFNKGKALEAISREKGTFLYGTPTWFVDILNPD 368
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 340 PANKEVKN-CAGVL 352
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 369 FSSYDISTMCGVI 382
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Result 14
US-09-997-653-194
; Sequence 194, Application US/09997653
; Publication No. US20030008297A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gettisen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavrin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann

```

```

; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Matanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC38
; CURRENT APPLICATION NUMBER: US/09/997, 653
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10

```

[illegible]

1	PRIOR FILING DATE: 1998-06-25
2	PRIOR APPLICATION NUMBER: 60/090678
3	PRIOR FILING DATE: 1998-06-25
4	PRIOR APPLICATION NUMBER: 60/090650
5	PRIOR FILING DATE: 1998-06-25
6	PRIOR APPLICATION NUMBER: 60/090694
7	PRIOR FILING DATE: 1998-06-25
8	PRIOR APPLICATION NUMBER: 60/090695
9	PRIOR FILING DATE: 1998-06-25
10	PRIOR APPLICATION NUMBER: 60/090696
11	PRIOR FILING DATE: 1998-06-25
12	PRIOR APPLICATION NUMBER: 60/090862
13	PRIOR FILING DATE: 1998-06-26
14	PRIOR APPLICATION NUMBER: 60/090863
15	PRIOR FILING DATE: 1998-06-26
16	PRIOR APPLICATION NUMBER: 60/091360
17	PRIOR FILING DATE: 1998-07-01
18	PRIOR APPLICATION NUMBER: 60/091478
19	PRIOR FILING DATE: 1998-07-02
20	PRIOR APPLICATION NUMBER: 60/091544
21	PRIOR FILING DATE: 1998-07-01
22	PRIOR APPLICATION NUMBER: 60/091519
23	PRIOR FILING DATE: 1998-07-02
24	PRIOR APPLICATION NUMBER: 60/091626
25	PRIOR FILING DATE: 1998-07-02
26	PRIOR APPLICATION NUMBER: 60/091633
27	PRIOR FILING DATE: 1998-07-02
28	PRIOR APPLICATION NUMBER: 60/091978
29	PRIOR FILING DATE: 1998-07-07
30	PRIOR APPLICATION NUMBER: 60/091982
31	PRIOR FILING DATE: 1998-07-07
32	PRIOR APPLICATION NUMBER: 60/092182
33	PRIOR FILING DATE: 1998-07-09

Query Match	5.4%;	Score 97;	DB 9;	Length 615;
Best Local Similarity	21.3%;	Pred. No. 2.3;		
Matches	54;	Conservative	50;	Mismatches 112;
				Indels 38;
				Gaps 10

QY	119	SNLSSPNTQOASSLLELAQKVGEGQIVYPRVASSSPTFGVLQOOLQSSSAVD-	177
Db	147	AGIILSVNPAFYQAMELEFLYKKGVCALVFPFQKFOQY--INVLKQICEPVENQPGA	204
QY	178	IKASDLPQSGDQVKNKDIQYKRTILDTSTVA--SQREYINSVKQKGPISNYVYGSESMCE	236
Db	205	LKSOFLPEDLTTIVSIVDAPELPGTLLEDEVVAAGSTROHLDOLQ-----YNQOFLS	255
QY	237	IKQIIRQOQVNOGLI-GTSGKPVYVT---DVLALNSLQLOEKQKV--AVEVINKLTLNT	289
Db	254	CHDPT-----NIQFTSGTSSPGKATLISHYNNINNSNIIIGERLKLHEKTEQOLRMILPNP	308
QY	290	LVIDLGLGLTEPANRKGIAHLAKSSNFYAQLSQOQPAKES-----RVRIECVD	339
Db	309	LYHCLGASVAGTMCMCMGATLILASPIFNKGKALLEAISRGFTGLYPTPMFVIDLQNPD	366
QY	340	PANKEVEN-CAGVL 352	
Db	369	FSYSDISTMCGVAI 382	

RESULT 15  
US-09-989-722-194  
Sequence 194, Application US/09989722  
Patent No. US20020072067A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Bornstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey  
 APPLICANT: Godowski, Paul J.  
 APPLICANT: Grimaldi, J. Christopher  
 APPLICANT: Gurney, Austin L.  
 APPLICANT: Kijavini, Ivar J.  
 APPLICANT: Napier, Mary A.  
 APPLICANT: Pan, James  
 APPLICANT: Paoni, Nicholas F.  
 APPLICANT: Roy, Margaret Ann  
 APPLICANT: Stewart, Timothy A.  
 APPLICANT: Tumas, Daniel  
 APPLICANT: Watanabe, Colin K.  
 APPLICANT: Williams, P. Mickey  
 APPLICANT: Wood, William I.  
 APPLICANT: Zhang, Zemin  
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 FILE REFERENCE: P2730P1C63  
 CURRENT APPLICATION NUMBER: US/09/989,722  
 PRIOR FILING DATE: 2001-11-19  
 PRIOR APPLICATION NUMBER: 60/049787  
 PRIOR FILING DATE: 1997-06-16  
 PRIOR APPLICATION NUMBER: 60/062250  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/065186  
 PRIOR FILING DATE: 1997-11-12  
 PRIOR APPLICATION NUMBER: 60/065311  
 PRIOR FILING DATE: 1997-11-13  
 PRIOR APPLICATION NUMBER: 60/066770  
 PRIOR FILING DATE: 1997-11-24  
 PRIOR APPLICATION NUMBER: 60/075945  
 PRIOR FILING DATE: 1998-02-25  
 PRIOR APPLICATION NUMBER: 60/078910  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/083322  
 PRIOR FILING DATE: 1998-04-28  
 PRIOR APPLICATION NUMBER: 60/084600  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/087106  
 PRIOR FILING DATE: 1998-05-28  
 PRIOR APPLICATION NUMBER: 60/087607  
 PRIOR FILING DATE: 1998-06-02  
 PRIOR APPLICATION NUMBER: 60/087609  
 PRIOR FILING DATE: 1998-06-02  
 PRIOR APPLICATION NUMBER: 60/087759  
 PRIOR FILING DATE: 1998-06-02  
 PRIOR APPLICATION NUMBER: 60/087827  
 PRIOR FILING DATE: 1998-06-03  
 PRIOR APPLICATION NUMBER: 60/088021  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088025  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088026  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088028  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088029  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088030  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088033  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088326  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088167  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088202  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088212  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088217  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088655

PRIOR FILING DATE: 1998-06-09  
 PRIOR APPLICATION NUMBER: 60/088734  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088738  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088742  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088810  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088824  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088826  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088858  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/088861  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/088876  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/089105  
 PRIOR FILING DATE: 1998-06-12  
 PRIOR APPLICATION NUMBER: 60/089440  
 PRIOR FILING DATE: 1998-06-16  
 PRIOR APPLICATION NUMBER: 60/089512  
 PRIOR FILING DATE: 1998-06-16  
 PRIOR APPLICATION NUMBER: 60/089514  
 PRIOR FILING DATE: 1998-06-16  
 PRIOR APPLICATION NUMBER: 60/089532  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089538  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089598  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089599  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089600  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089653  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089801  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089907  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089908  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089947  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/089948  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/089952  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/090246  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090252  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090254  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090349  
 PRIOR FILING DATE: 1998-06-23  
 PRIOR APPLICATION NUMBER: 60/090355  
 PRIOR FILING DATE: 1998-06-23  
 PRIOR APPLICATION NUMBER: 60/090429  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090431  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090435  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090444  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090445  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090472  
 PRIOR FILING DATE: 1998-06-24

```
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09
```

Query Match 5.4%; Score 97; DB 10; Length 615;  
Best Local Similarity 21.3%; Pred. No. 2.3;  
Matches 54; Conservative 50; Mismatches 112; Indels 38; Gaps 10;

```
119 SNFLSSPNGTQOASLLEAKVYEQIYVPDVASSSFTVFGLYOQLLOSSSAVD- 177
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
147 AGILIVSNPAVQAMELEVVVKVGCALVFPKOPKTOY--YNYLKOICEVENAOPGA 204
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 178 IKASDLPQSGQOVNKDITQKRTILDSFTVA-SOREYINSVKQKPISNYYVYESMCE 236
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 205 LKSOQLPDLTVISVDAPLPGLILDEVAAGSTRQHLDLO-----YNOQLUS 253
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 237 IKDIIIDQYVNVOLI-GTSDKPYVYT---DVLALNSNLCDEKQKV---AVEVIKNLNTT 289
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 254 CHDPF-----NIQFTSGTGSPEKATLSHYNI VNNNSNITGERLKLHEKTPBQLRWILPNP 308
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 290 LVLDLIGLGLIPANKNGIAHLAKSNFYAQLSQQFDKASE-----VRILRCVD 339
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 309 LYHCGSVAGTMMCLMYATLILASPIFNKKALFAISRERGTFLYGTPTMFVDILNOPD 368
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 340 FANKEVKN-CAGVL 352
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 369 FSSYDISTMCGGVI 382
```

Search completed: January 24, 2003, 19:43:54  
Job time : 17 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2003, 19:40:55 ; Search time 17 Seconds  
(without alignments)  
616.151 Million cell updates/sec

Title: US-09-675-509-4  
Perfect score: 1802  
Sequence: 1 MSTOPTLTVGLFPYLPSPWN.....CVDPAKRVKAGVLRPFL 356

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

7 number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	139.5	7.7	10182	US-09-134-001C-3159	Sequence 3159, Ap
2	103.5	5.7	397	US-09-134-001C-4104	Sequence 4104, Ap
3	98.5	5.5	1149	US-07-915-203-2	Sequence 2, Appl
4	98.5	5.5	1149	US-08-272-887-2	Sequence 2, Appl
5	98.5	5.5	1149	US-08-769-449-2	Sequence 2, Appl
6	97.5	5.4	624	US-09-336-447A-7	Sequence 7, Appl
7	97.5	5.4	639	US-09-134-001C-5661	Sequence 5661, Ap
8	97.5	5.4	825	US-09-540-824-26	Sequence 26, Appl
9	97.5	5.4	2037	US-09-306-998-3	Sequence 3, Appl
10	93	5.2	348	US-07-989-845-2	Sequence 2, Appl
11	93	5.2	348	PCT-US93-11298-2	Sequence 2, Appl
12	92.5	5.1	607	US-09-204-208A-11	Sequence 11, Appl
13	92.5	5.1	831	US-09-269-861A-8	Sequence 8, Appl
14	92.5	5.1	3418	US-08-639-501-2	Sequence 2, Appl
15	92.5	5.1	3418	US-08-603-753D-4	Sequence 4, Appl
16	92.5	5.1	3418	US-09-044-946-2	Sequence 2, Appl
17	92.5	5.1	3418	US-08-755-587-44	Sequence 44, Appl
18	92.5	5.1	3418	US-09-044-908-2	Sequence 2, Appl
19	92.5	5.1	3418	US-09-099-753-4	Sequence 4, Appl
20	92.5	5.1	3418	US-08-986-106-4	Sequence 4, Appl
21	92	5.1	468	US-08-961-083-10	Sequence 10, Appl
22	91.5	5.1	386	US-09-134-001C-3809	Sequence 3809, Ap
23	91.5	5.1	1211	US-09-134-001C-4820	Sequence 4820, Ap
24	91	5.0	279	US-09-134-001C-4222	Sequence 4222, Ap
25	91	5.0	1285	US-07-582-945-2	Sequence 2, Appl
26	91	5.0	1285	US-08-453-141-2	Sequence 2, Appl
27	91	5.0	1285	US-08-293-314-2	Sequence 2, Appl

28	90.5	5.0	707	US-09-134-001C-2962	Sequence 2962, Ap
29	90.5	5.0	2366	US-08-480-604A-10	Sequence 10, Appl
30	90.5	5.0	2366	US-08-405-496A-10	Sequence 10, Appl
31	90.5	5.0	2366	US-08-915-136-10	Sequence 10, Appl
32	90.5	5.0	2366	US-08-957-310-10	Sequence 10, Appl
33	90	5.0	384	US-08-220-958-2	Sequence 2, Appl
34	90	5.0	434	US-09-134-001C-4251	Sequence 4251, Ap
35	90	5.0	454	US-09-416-050A-6	Sequence 6, Appl
36	90	5.0	454	US-09-664-800-6	Sequence 6, Appl
37	90	5.0	454	US-09-665-309-6	Sequence 6, Appl
38	90	5.0	454	US-09-661-569-6	Sequence 6, Appl
39	90	5.0	1024	US-09-091-117-5	Sequence 5, Appl
40	89.5	5.0	1098	US-07-777-715-7	Sequence 7, Appl
41	89.5	5.0	1098	US-08-170-126-2	Sequence 2, Appl
42	89.5	5.0	1098	US-08-954-418-2	Sequence 2, Appl
43	89.5	5.0	2183	US-08-348-891A-7	Sequence 7, Appl
44	89.5	5.0	2183	US-08-905-817-7	Sequence 7, Appl
45	89	4.9	892	US-09-336-447A-5	Sequence 5, Appl

## ALIGNMENTS

RESULT 1  
US-09-134-001C-3159  
; Sequence 3159, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/134, 001C  
; PRIOR FILING DATE: 1998-08-13, 001C  
; PRIOR APPLICATION NUMBER: US 60/064, 964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055, 779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 3159  
; LENGTH: 10182  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3159

Query Match 7.7%; Score 139.5; DB 4; Length 10182;

Best Local Similarity 20.5%; Pred. No. 0.0095;

Matches 89; Conservative 59; Mismatches 140; Indels 147; Gaps 17;

QY	24	NEVKILNLIKIDPTQVSGYNIETFEPCYSDASLQSLPVFSFDSIFLPLVSLGKVS	83
DB	8873	NFTQAINNARDAL-NKTOGGQNLDFNAIDTFKD-----DIFKTKD-----ALNGIER	8917
QY	84	L-----DESLVRGVTGDLHSFVSSASVNGSV-----	110
DB	8918	LTAKSAKAEKLIDSLKINKAQFTHANDEIMNTNSINQSLRIYQARDLNDAMKSLDEL	8977
QY	111	--YGPQYLSNPLSSPNGTQOASSLLELAQV-----GYEQIVYDVA	153
DB	8978	NNQAFPPQOASNYNSINSEDLKQCFDHALSNARKVLAKENGNLDEKOIOGKQVY-----	9032
QY	154	SSSFTVFGLYO-----QLLOS-----SSSAVDIKASD	182
DB	9033	EDTDALNGIORLSKAKAKAIQYQOSLSYINDQRIHAENNINNSDLSTLANTLSKASD	9092
QY	183	LPGSDQVNDITQKYRTIILDSVVASQREYINSVKQKPISNYYGYSBSMCKIDIR	242
DB	9093	L---DNAMDLRPTIS--NSTSVPSNVYINADK-----NLQIEFDEAL-----	9132
QY	243	DQYINVOLIGTSKRPYYTVDVALNSNLCDK-----QYVAEVIYKLLTNTLVLDLG	296
DB	9133	-QOASATSSSTSENPAITEVLGLSQAIYDTKNALNCEQRLATEKSKDLTKLIGKLDKLNK	9191

QY	297	LGTLTPAK---	NCIAHLAKSNFAYQLQOPDAKESERV---	-----RCVDF--	340
Db	9192	AQLEDVNTKVASANTLTETLSQLTGSTLELNDKMKLRDKLTLVNPVKASLNVNRMDYNL			9251
QY	341	---ANKEVKNCAGVL			352
Db	9252	KRQFKALKLEAKGVL			9266

```

RESULT 2
US-09-134-001C-4104
; Sequence 4104, Application US//09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4104
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-4104

```

```

Query Match Similarity 5.7%; Score 103.5; DB 4; Length 397;
Best Local Similarity 18.5%; Pred. No. 0.14;
Matches 74; Conservative 72; Mismatches 142; Indels 111; Gaps 18;

QY 52 CYSDASLQSLPDVF--STDSTFLPYLVLGQ-VKSLDE-----SLVRGV 92
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1 CYNDINIIENNEVEFAMKSFTEFFMHNIVAMGCVKSVTOLANTLAKEGHPVTIISVRGA 60
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 93 TG---DLHSFV-----SSASVNSVYGFPGYLCNFISSNPGNQQASSL 136
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 61 DSPYELHSAIVKVVVDYRLKLTNRATIANRIKKYTPPLNTKVLISQEPFGSQSSSY 120
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 137 ELAQKVGEGQIYYPD---VASSSFTVFGLYQOLLQSSSA---AVDIKASDLPQSGDQ 189
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 121 E---KKMIKAIKHTKTDVLVGRASPNI-----LISKAKAEIVITIAEMHMFADHPDQ 171
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 190 VAKDITQKRTT--LDSTVASQREY-----INSVKQKPIG---- 224
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 172 YQKEIIAAYRNINKITTLTVADQOKYQSLKTPVVVIPMMVTEKRIAPAKKNIISAGRL 231
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 225 NYVYVGS---ESMCEIKDIRDOQVNVOLIGTSDKP-----VYVYDVLA 265
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 232 EYEKQYDILBESIRLIQELROLANDVHIIYGGSSKTSLVDFINQYHLNDLKIYEPTQE 291
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 266 LNSNLCDK-----QKVAVEVIKNLLTNTLVLDLG-LGTLTPANKGIAHLKSSN 316
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 292 LNNKLAQSKIVVPSRNEGFGMIILEAMQDNIVISFEGNVGPDSTINNGDNGYLVNVEY 351
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 317 FYAQLSQGFDAKESVRVLCYDFPANKVEYKNCAGVLRPF 355
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 352 -VSELAKRIDLTTOHYNELDHI-----IENSKDTLKQF 383
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

RESULT 3  
US-07-915-203-2

```

APPLICANT: Ogiwara, Katsutoshi
APPLICANT: Sakanaka, Kazunobu
APPLICANT: Hori, Hidetaki
APPLICANT: Asano, Shouji
APPLICANT: Kawasugi, Tadaaki
TITLE OF INVENTION: No. 5359048el Microorganism and Insecticide
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Releasee #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/915,203
FILING DATE: 19920723
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: M/K 301
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1149 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein

```

Query Match	5.5%;	Score 98.5;	DB 1;	Length 1149;
Best Local Similarity	22.6%;	Pred. No. 2.3;		
Matches	91;	Conservative	58;	Mismatches 126; Indels 127; Gaps 25;
QY	26	VKLINLIKDVLPFOVSGYN-----IEYEFD-CYSADASIQLPDPVSTDSIFLP-YLV	76	
Db	728	VQARNLQDQGFNRINGENENMGWGTGTEIEVEVGDLVFDKRSIR-LTSAREIDTETPTVL-	785	
QY	77	SLGVSRLDSISVRGVG-DLHSPVSSASV-----NGSYGPP-----	114	
Db	786	---YQIDESLLEPYRYRYLTKGFTGSSQDIEIKLIRHRAQIVKVPDNLIPDRPNS	841	
QY	115	-----QYLCSNFL--SSPNGTQOAS-----LLELAQKVGEQIYVDPV	152	
Db	842	CGGVDRCSGGQYDANLALENGENGMSSDHAFSFHDTGEIDLMENTGI-MIVFKIP	900	
QY	153	ASSSSFFVFGIYQOQLLOSSSAVDIASDLPQSGDOVN--KDTQKRYRTLLDSTVASQ	210	
Db	901	TTNGNAATLGNL--EFVEEG-----PLSGETLEMAQOQEQOQDMARKARASE	946	
QY	211	REYINYSKQGGKPISNYYVGYSE---SMCEIKDIIIRQOQYVVOGLIGTSDKREYVYTDVL-	264	
Db	947	KTYI-AAKQA-IDRLPADYQDQKLNNGVEMSDLLAQN-L-VQGI-----PYVNDALPE	997	
QY	265	--ALNSNLCEKQKQKAVAVIKNLLTNTL--VLDLIGLGLTPAN--KNGIAHLAKSS--	315	
Db	998	IPGMNYSFTPE-----LTNLLQOAMNLYDLQNALPNQDPFRNGLSNMNATSDVN	1045	
QY	316	-----NFAQLSQQPDAKSESVRYLACVFPANKE	344	
Db	1046	VQQLSDTSVLVIPMNSQVSOQFTVQGNRYRYLVVT--ARRE	1085	

RESULT 4  
US-08-272-887-2

```

: Sequence 2, Application US/08272887
: Patent No. 5747450
: GENERAL INFORMATION:
: APPLICANT: Ohba, Michio
: APPLICANT: Iwahana, Hidenori
: APPLICANT: Sato, Reichi
: APPLICANT: Suzuki, No. 5747450ukazu
: APPLICANT: Ogiwara, Katsutoshi
: APPLICANT: Sakana, Kazunobu
: APPLICANT: Hori, Hideaki
: APPLICANT: Amano, Shouji
: TITLE OF INVENTION: No. 5747450e1 Microorganism and Insecticide
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: David R. Saliwanchik
: STREET: 2421 N.W. 41st Street, Suite A-1
: CITY: Gainesville
: STATE: FL
: COUNTRY: US
: ZIP: 32606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/272,887
: FILING DATE: 08-JUL-1994
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/915,203
: FILING DATE: 23-JUL-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Saliwanchik, David R.
: REGISTRATION NUMBER: 31,794
: REFERENCE/DOCKET NUMBER: M/K 301
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 904-375-8100
: TELEFAX: 904-372-5800
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1149 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-272-887-2

```

[illegible][illegible]

```

RESULT 5
US-08-789-449-2
; Sequence 2, Application US/08789449
; Patent No. 5824878
; GENERAL INFORMATION:
; APPLICANT: Ohba, Michio
; APPLICANT: Iwahana, Hidemori
; APPLICANT: Sato, Reichi
; APPLICANT: Suzuki, No. 5824878ukazu
; APPLICANT: Ogiwara, Katsutoshi
; APPLICANT: Sakanaka, Kazunobu
; APPLICANT: Hori, Hidetaki
; APPLICANT: Asano, Shouji
; APPLICANT: Kawasugi, Tadaaki
; TITLE OF INVENTION: No. 5824878el Microorganism and Insecticide
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/789,449
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/272,887
; FILING DATE: 08-JUL-1994
; APPLICATION NUMBER: US 07/915,203
; FILING DATE: 23-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: M/K 301
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1149 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-789-449-2

Query Match 5.5%; Score 98.5; DB 2; Length 1149;
Best Local Similarity 22.6%; Pred. No. 2.3;
Matches 91; Conservative 58; Mismatches 126; Indels 127; Gaps 25;

QY 26 VKLILNIDVLPQVSGN-----ITYEFD-CYSDASLQSLPDVSTDSIFLP-VLV 76
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 728 VQARVLDDTGFRNIRNGENGTGTGIEVVGCVLFLKORSLR-LTSAEIDITETVPTL- 785
QY 77 SLGGSVLSDESLVRCVGT-DLHSFVSSASV-----NGSVYGF----- 114
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 786 ----YQVIDESLTKRYTTLKLGKFGIGSQDDEIKLIRHRANQIVGNVDDNLIPVPRVNS 841
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 115 -----QYLCGNFLL--SSPNGTQOASS-----LLELAQKVGEOIVPYDPV 152

```

```

Db 842 CGVDRCEQOQYVDANLALENNGENGNMSSDBSHAFSHIDGELDINENTGI-WLVEKIP 900
QY 153 ASSSFTVFGLYQOQLSSSSAAVIDIKASDLPQSGDQVN-KDITQKYRTTIDSTFVVASQ 210
Db 901 TTNGATNATIGNL-EFVEEG-----PLSGETLEMAQOQBQOMODKMARRRKAAASE 946
QY 211 REYNISVQGRKISNYVYGSE-----SMCEIDIIIRDQOYNVOLIGTSKPYVYTDVL-- 264
Db 947 KTYV-AAQOA-IDRLFADYQDQKLSGVEMSDLLAQN-L-VQSI-----PYVNDALPE 997
QY 265 --ALNSNLCDEKQKVAVEYIKNLTTNTL--VDLLGLGLTLPAN-KNGIAHLAKSS--- 315
Db 998 IFGNNYTSFTE-----LNNRLQAMNLYVDLQNAIRNGDPFRNGLSMWNATSDVN 1045
QY 316 -----NFAQLSQQFDAKESEVRVLRCDVPFANKS 344
Db 1046 VQQLSDTSVLYVPMNNSQVSGQFVQPNRYVLRVT--ARKE 1085

RESULT 6
US-09-336-447A-7
; Sequence 7, Application US/09336447A
; Patent No. 6310190
; GENERAL INFORMATION:
; APPLICANT: HANSEN, ERIC J.
; APPLICANT: AEBI, CHRISTOPH
; APPLICANT: COPE, LESLIE D.
; APPLICANT: MACIVER, ISOBEL
; APPLICANT: FISKE, MICHAEL J.
; APPLICANT: FREDENBURG, ROSS A.
; TITLE OF INVENTION: USPA1 AND USPA2 ANTIGENS OF MORAXELLA CATARRHALIS
; FILE REFERENCE: AMCY:024
; CURRENT APPLICATION NUMBER: US/09/336,447A
; CURRENT FILING DATE: 1999-06-21
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-09-336-447A-7

Query Match 5.4%; Score 97.5; DB 4; Length 624;
Best Local Similarity 25.4%; Pred. No.1.1;
Matches 67; Conservative 47; Mismatches 89; Indels 61; Gaps 18;

QY 102 SSASVNGSVYGGFPXYLCSNF--LSSPNGTQOASSLLELAQKGYE-QIVY---PDVASS 155
Db 230 SGARNG-----LTKNFPVSGVSRADNQTIVSGS-----QKGRQIVHGAQISPD 275
QY 156 SSFTVFG--LYOQLLOSSSSAAVIDIKASDLPQSGDQVNKDITQKYRTTIDSTFVVASQ 213
Db 276 STDVANGSOLY-----ALATAVDNNOYDIEINQDNI-KDL-QKEVKGDKREVGLSRD- 326
QY 214 INSVAQGRKISNYVYGSSMCEIKDIIIRDQOYNVOLIGTSKPYVYTDVLALNSNLDE 273
Db 327 IGSLLHD--DVAQNOADIANKKADIKEL--DKEMNV-----LSRIVSLINDVADN 372
QY 274 KQKVA-----VEVINKLITNTLVLDLGLTLPAN-KNGIAHLAKSSNFAQLSQQFDA 327
Db 373 QADIAKNQADITLERNVEGL-LDSGLLIDQKADIDNNINHI-----YELAQOQDO 424
QY 328 KESEVRVLRCDVPFANKS--VKKCA 349
Db 425 HSDIKTLAKASAAVTDRIAKKKA 448

```

[illegible]

QY 79 GGVKSLDESLVGVGTGDLHSFVSSSASVNGSVYGFPOYLCSNFISSPNTQOASSLLEL 138  
Db 467 TROPIMISELLEHLEHLSIAQO-----FPSALGISF-ISVVEGWRK-----RL 510  
QY 139 AOKVGYQIYPPDVASSSFTVFGLYQOQLLOSSSSAAVDK-----ASDLPQSGDVNK 192  
Db 511 AKSVYFPEIKPEI---SDLLFFNLGTSIFPTSDKKHIVSPVMTMAESLOSPPADSL 567  
QY 193 DITOK-----YRTILDSTVVASOREVI-----NSVKOGKP 222  
Db 568 DVCKKLYIALFLKFSYSHRYVEVTAVSQALYLYPNFISIVPGTALPBLKE--- 624  
QY 223 ISNVYVSSSKEIKDIIRDQYVQLIGTSKRP---YVYTDVALNSNCDE-----K 274  
Db 625 -----KONLPAIQDISDEPQRLSLYELEELPTGLLOSSILFTLNLIMAMIDYFK 676  
QY 275 QKXAVEI-----KULLNTLVLDLIG----- 296  
Db 677 EQAFIEIFVPIMDMLQYSLKELSLKRLSEKLSLTLQAVSDSIESAKANKRKLALQSHR 736  
QY 297 -LGLT--LPANKNGIAHLAKSSNFY-----AQL-----SQGFPAKSEVRLR 336  
Db 737 PLGITSQVPFEEGYS-LDKSSHIDPERAQLNKLRAQHRDAKKGAIRTLR 786

RESULT 9  
US-09-306-998-3  
Sequence 3, Application US/09306998  
Patent No. 6291173  
GENERAL INFORMATION:  
APPLICANT: Bartel, Paul L.  
APPLICANT: Tavitgian, Sean V.  
TITLE OF INVENTION: MMSC2- An MMAC1 Interacting Protein  
FILE REFERENCE: MMSC2  
CURRENT APPLICATION NUMBER: US/09/306,998  
CURRENT FILING DATE: 1999-05-07  
EARLIER APPLICATION NUMBER: US 60/084,740  
EARLIER FILING DATE: 1998-05-08  
NUMBER OF SEQ ID NOS: 72  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 2037  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-306-998-3

Query Match 5.4%; Score 97.5; DB 4; Length 2037;  
Best Local Similarity 21.4%; Pred. No. 7.2;  
Matches 75; Conservative 48; Mismatches 147; Indels 81; Gaps 15;  
QY 7 TLTAVGLFPVYLPSPNNGENVEKILNLRKDVLPYQVSGVNIETEDCYSDASLQSPVFS 66  
Db 778 TVAVIGVAKPLPLSPEDC---YVSAKED-----SFLYPPHSC-EEAGLADKP-LFR 822  
QY 67 TDSIFLPYVLSLGVKSLDESLVGVGTGDLHSFVSSASVNGSVYGFPOYLCSNFISSP 126  
Db 823 ADLA-----LVGTNDADVLDS-----TFESPSPENDSITYSQAAILSHGSSCG 868  
QY 127 NGTQOASSLLELAOKVGYEQIYPPDVASSSFTVFGLY-----QQLLOSSSSAAVDI 178  
Db 869 DGNVYSSLPSPSPK-----DVTENSCDPEVLDLHMSLEELYTQNLQORDENTPSV 919  
QY 179 KASDLPQSGDVN-----KQITOKYRTILDSTVVASOREVINSVKOKPISNVYVGSSES 233  
Db 920 DISMGPASGTTINDYPPANAIEQYF--CENTTIVWTESHPSEV-----IS 963  
QY 234 MCEIKDIIRDQYVQLIGTSKRPVYVTDVALNSNLCDEKQKXAVEVIRKLLTNTLVLD 293  
Db 964 SAEIPLVLPDS-----AGKSEVLEQSSSLACNAE-CVWLQVNSKSEFR--TINIAGK 1014  
QY 294 LGLGLTLPLANKNGIAHLAKSSNFYQOLSQGFPAKSEVRLRCVDPANKE 344  
Db 1015 NSSLGMTVSANKGIGLVRSIIHGAIIS-----RDGRIRAGCIIISINEE 1060

RESULT 10  
US-07-989-845-2  
Sequence 2, Application US/07989845  
Patent No. 5304472  
GENERAL INFORMATION:  
APPLICANT: Baes, Steven  
APPLICANT: Swartz, James  
TITLE OF INVENTION: METHOD OF CONTROLLING POLYPEPTIDE  
TITLE OF INVENTION: PRODUCTION IN BACTERIAL CELLS  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080-4990  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/989,845  
FILING DATE: 19921120  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Haack, Janet E.  
REGISTRATION NUMBER: 28,616  
REFERENCE/DOCKET NUMBER: 752  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1896  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 348 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
US-07-989-845-2

Query Match 5.2%; Score 93; DB 1; Length 348;  
Best Local Similarity 23.5%; Pred. No. 1.2;  
Matches 52; Conservative 42; Mismatches 79; Indels 48; Gaps 10;  
QY 100 VSSASAVNGSVYGFPOYLCSNFISSPNTQOASSLLELAOKVGYQIYPPDVASSSFT 159  
Db 23 VFAEASLTGAGATFPAPYAKWA-----DYOK-----ETGNKKNYOGI-----GSS-- 64  
QY 160 VFGLYQQLLOSSSSAAVDIKASDLPQSGDVNKTQKRTILDSVVASOREVINS--- 216  
Db 65 --GGVQIITANT---VDGASDAPLSDEKLAQEGLFQFTYIGGVVLANIPGLSGEL 118  
QY 217 VKQGRISNVYGY-----SESNCEIKDIIRDQYVQLIGTSD---KPYVYTDVALNS 268  
Db 119 VLDGKTLGIVIGIKIKMDDEAIKLNPGKLPQGVIAVVRADGSGTSFVFSYLA--- 175  
QY 269 NLCEKQKXAVEVIRKLLTNTLVLDLGLGLTLPLANKNGIA 309  
Db 176 -----KYNEMKNNVGTGTYKWPILG---GKGNDGIA 206  
RESULT 11  
PCT-US93-11298-2  
Sequence 2, Application PC/TUS9311298  
GENERAL INFORMATION:  
APPLICANT: Genentech, Inc.  
TITLE OF INVENTION: METHOD OF CONTROLLING POLYPEPTIDE PRODUCTION IN

TITLE OF INVENTION: BACTERIA  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/11298  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Hasak, Janet E.  
REGISTRATION NUMBER: 28,616  
REFERENCE/DOCKET NUMBER: 752  
TELEPHONE: 415/225-1896  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 348 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
PCT-US93-11298-2

Query Match 5.2%; Score 93; DB 5; Length 348;  
Best Local Similarity 23.5%; Pred. No. 1.2;  
Matches 52; Conservative 42; Mismatches 79; Indels 48; Gaps 10;

QY 100 VSSASVNGSYGPPQYICSNFLSSPNGTQOASLLELAQKVGEOIVYDVASSSFT 159  
DB 23 VFAEASLIGAGATPPAPYAKWA-----DTYQK-----ETGNKVNQGI-----GSS-- 64  
QY 160 VFGLYQQLQSSSSAAVDIKASDLPQSGDQVKNKDTOKYRTLLDSTVVASQREYINSV-- 216  
DB 65 --GGVKQITANT---VDFGASDAPLSDSEKLAQEGLPFPVTVIGSVLAVNIPGKSGEL 118  
QY 217 VKQKPISNYYVG-----SESMCEIKDIIIRDQYNVOLIGTSD---KPYVYTVLALNS 268  
DB 119 VLDEGTLLDILGKIKKMDDEIAIKLNGKLPQGNIAVVRADSGSGTSFVTSYLA--- 175  
QY 269 NLDEKQYAAVEVINKLLTNTVLVDLGLGLTPANKKGIA 309  
DB 176 -----KYNEEMKNVGVSTVKWPIGIG--GKANDGIA 206  
RESULT 12  
US-09-204-208A-11  
Sequence 11, Application US/09204208A  
Patent No. 6399320  
GENERAL INFORMATION:  
APPLICANT: Markau, Ureula  
APPLICANT: Ebenbichler, Christine  
APPLICANT: Achhammer, Gunthar  
APPLICANT: Ankenbauer, Waltertraud  
TITLE OF INVENTION: Modified DNA-Polymerase from Carboxydotherrnus hydrogenofornans at  
TITLE OF INVENTION: use for Coupled Reverse Transcription and Polymerase Chain React  
FILE REFERENCE: 4765  
CURRENT APPLICATION NUMBER: US/09/204,208A  
CURRENT FILING DATE: 1998-12-01  
PRIOR APPLICATION NUMBER: EP/97121151.1  
PRIOR FILING DATE: 1997-12-02

NUMBER OF SEQ ID NOS: 12  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 11  
LENGTH: 607  
TYPE: PRT  
ORGANISM: Carboxydotherrnus hydrogenofornans  
US-09-204-208A-11

Query Match 5.1%; Score 92.5; DB 4; Length 607;  
Best Local Similarity 19.5%; Pred. No. 3.1;  
Matches 67; Conservative 52; Mismatches 110; Indels 115; Gaps 15;

QY 21 ENGENEKLNLIDVLPVQSGVNIETEDDCYSASL-----QSLPVDV-----FST 67  
DB 237 ELGEQIE--ELIREI-----VVLAGEEFNLSPROLGYILFEKGLPIYKTKTKGYST 287  
QY 68 DS-----IFLPYLVSLGQVKSLE-----SLVGVGTDLHSFVSSASVNGSV 110  
DB 288 DAELYEELLPFHETIGKILNRYQMLKSTYTDGLMPLINERTGKLTHTTNGTGLGR- 346  
QY 111 YGFPQYLCNFFLSSPNGTQOASLLELAQV-----GYEQIYPDVASSSFTVFG 162  
DB 347 -----LASSERNINIPRIELGRKRMFIPSPGYDIVSAD----- 384  
QY 163 LYQQL-LQSSSSAAVDIKASDLPQSGDQVKNKDTOKYRTLLDSTVVASQREYINSV-- 218  
DB 385 -YSQIEIRLHLHFSEEPKLEAVQKGEDIRKTASEVFGVSLPEVPEMRAHAKSVFGI 443  
QY 219 -----QKPISNYYVG-----SESMCEIKDIIIRDQYNVOLIGTSDK 256  
DB 444 VYGISDPLGRDLKIPREYAGKTIKXIFANYPKRYEVLDELVTAREKGVTTLLFGR--R 501  
QY 257 PYV-----YTDVLAALNSNLCEDEKQYAAVEVINKLLTN 288  
DB 502 RYIPELSSKNRTVQGFGEFTAMNTPL-----QGSAAIDIIKLAMIN 541

RESULT 13  
US-09-269-861A-8  
Sequence 8, Application US/09269861A  
Patent No. 6468775  
GENERAL INFORMATION:  
APPLICANT: Ankenbauer, Waltertraud  
APPLICANT: Markau, Ureula  
APPLICANT: Svetlichny, Vitaly  
APPLICANT: Schmitz-Agheguyan, Gudrun  
APPLICANT: Reiser, Astrid  
APPLICANT: Angerer, Bernhard  
APPLICANT: Laue, Frank  
APPLICANT: Bonch-Osmolovskaya, Elizaveta  
TITLE OF INVENTION: THERMOSTABLE DNA POLYMERASE FROM CARBOXYDOTHERMUS HYDROGENOFORNAN  
FILE REFERENCE: 4494  
CURRENT APPLICATION NUMBER: US/09/269,861A  
CURRENT FILING DATE: 1999-11-22  
PRIOR APPLICATION NUMBER: PCT/EP97/05391  
PRIOR FILING DATE: 1997-10-01  
PRIOR APPLICATION NUMBER: EP/96115873.0  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 8  
LENGTH: 831  
TYPE: PRT  
ORGANISM: Carboxydotherrnus hydrogenofornans  
US-09-269-861A-8

Query Match 5.1%; Score 92.5; DB 4; Length 831;  
Best Local Similarity 19.5%; Pred. No. 5.1;  
Matches 67; Conservative 52; Mismatches 110; Indels 115; Gaps 15;  
QY 21 ENGENEKLNLIDVLPVQSGVNIETEDDCYSASL-----QSLPVDV-----FST 67

Db 461 ELAGEQIE-ELIREI-----YVLAGEEFPNLSPROGLVLEFKGLPVIKKTIGYST 511  
 Qy 68 DSIFLPLYL-----VSLGCVKSLDE-----SLVGVGTGDLSPFSSASVNGSV 110  
 Db 512 DAEVLEELPFPHEIGIKLIINYROMLKSTYDMLPLINERCKLMTTNGTTLGR- 570  
 Qy 111 YGFPQYLCNFFLSSPNGTQOASSLLELAQV-----GYEQIYVPDVASSSFTVFG 162  
 Db 571 -----LASSEPNLQNIPIRELEGRKLRKMFIPSGDYDIVAD----- 608  
 Qy 163 LYQOL-QOSSSSAAVDIKASDLPQSGDVNKDITQKRTITLDSVVASQREYINSVK--- 218  
 Db 609 -YSQIELRLAHFSEEEKLLEAVQKGEDIRKTASEVFGVLEEVTPDMAHAKSVNFGI 667  
 Qy 219 -----OGKPISNYYVG-----SESMCEIKDIIRDOQYVNOIGTSDK 256  
 Db 668 VGISDQGLGRDLKIPREVAQKTIKYNFANYPKREYLDLVRFAREKGYTTLFG-R 725  
 Qy 257 PYV-----YTDVLALNSLCEKQKQKVAEVIKMLLTN 288  
 Db 726 RYIPELSSKRTVQGFERTAMNTPL-----QGSADIIKLAMIN 765

RESULT 14  
 US-08-639-501-2  
 Sequence 2, Application US/08639501  
 Patent No. 5837492

GENERAL INFORMATION:  
 APPLICANT: Tavcigian, Sean V.  
 APPLICANT: Kamb, Alexander  
 APPLICANT: Simard, Jacques  
 APPLICANT: Couch, Fergus  
 APPLICANT: Rommens, Johanna  
 APPLICANT: Weber, Barbara  
 TITLE OF INVENTION: Chromosome 13-Linked Breast Cancer  
 TITLE OF INVENTION: Susceptibility Gene  
 NUMBER OF SEQUENCES: 124  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Venable, Baetjer, Howard & Civiletti  
 STREET: 1201 New York Avenue N.W., Suite 1001  
 CITY: Washington  
 STATE: DC  
 COUNTRY: USA  
 ZIP: 22204

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/639,501  
 FILING DATE: 29-APR-1996  
 CLASSIFICATION: 530  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/585,391  
 FILING DATE: 11-JAN-1996  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/576,559  
 FILING DATE: 21-DEC-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/575,359  
 FILING DATE: 20-DEC-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/573,779  
 FILING DATE: 18-DEC-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Ihnen, Jeffrey L.  
 REGISTRATION NUMBER: 28,957  
 REFERENCE/DOCKET NUMBER: 24884-116802-04  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-962-4810  
 TELEFAX: 202-962-8300  
 INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 3418 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-639-501-2

Query Match 5.1%; Score 92.5; DB 2; Length 3418;  
 Best Local Similarity 21.7%; Pred. No. 49;  
 Matches 86; Conservative 55; Mismatches 152; Indels 103; Gaps 22;

Qy 22 NGNEVKILNLIKD-----VLPTQVSGYNIETED---CYSDASLQSLPDVFSTDS 69  
 Db 270 SGNSFFK-VNSCKDHIGKSMNVLEDEYETVVDTSSEDSFSLCFSKCRTNLOKVRTSKT 328  
 Qy 70 ---IFLPLYLVSIGVSLDE-----SLVGV-----TGDLHS-----FVSS 103  
 Db 329 RKKIF-----HEAVADCEKSKQVKEKYSFVSEVENDTDPDLSNVAHQKPFESGS 380  
 Qy 104 ASVNGSVYGFPOYL-C-SNFLSSPNGTQ-QASSLELA---QKVGEOIYVDPVASSSS 157  
 Db 381 DKISKEV--VPSLACEWSQTLTSLGNGAQWMEKIPLHISCDQNISKDLDTENKRRKD 438  
 Qy 158 FTVFGLYOOLLQSSSSAAVDIKASDLPQSGDVNKDITQKRTITLDSVVASQREYINSV 217  
 Db 439 F-----LTSNSLIP--RISSLPKSEKPLNBEVNVNKR--EBOHLSHTDCLAV 484  
 Qy 218 KQG-----KPISNYYVGSESMCEIKDIIRDOQYVNOIGTSDKPYVYTDVLALNSNL--- 270  
 Db 485 KQAIISTGPSVASSFQGIKSIPIRISPK-ETNASFSGMTDPNFKKETEASESGEIH 543  
 Qy 271 ---CDEKO-KVAEVIKN-----LNTVTLVDLGLGLTPANKNGIAHLAKSNFY--- 318  
 Db 544 TVCSQKEDSLCPNLIDNGSPATTQNSVALKNAGLSTLKKTKNTKIYAIHDETFYKKG 603  
 Qy 319 -----AQLSQFDASEVRLRCVDPANKE 344  
 Db 604 KIPDKQKSELINCSAQFEANAFAP---LTFANAD 635

RESULT 15  
 US-08-603-753D-4  
 Sequence 4, Application US/08603753D  
 Patent No. 5891857

GENERAL INFORMATION:  
 APPLICANT: Holt, Jeffrey T.  
 APPLICANT: JENSEN, ROY A.  
 APPLICANT: PAGE, DAVID L.  
 APPLICANT: KING, MARY-CLAIRE  
 APPLICANT: SZABO, CSILLA I.  
 APPLICANT: JETTON, THOMAS L.  
 APPLICANT: ROBINSON-BENION, CHERYL L.  
 TITLE OF INVENTION: CHARACTERIZED BRCA1 AND BRCA2  
 TITLE OF INVENTION: PROTEINS AND SCREENING AND THERAPEUTIC METHODS BASED ON  
 NUMBER OF SEQUENCES: 29  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: ARLES A. TAYLOR, JR.  
 STREET: SUITE 1401, UNIVERSITY TOWER, 3100 TOWER  
 CITY: DURHAM  
 STATE: NORTH CAROLINA  
 COUNTRY: USA  
 ZIP: 27707

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 800 KB storage  
 COMPUTER: IBM PC/XT/AT compatible  
 OPERATING SYSTEM: Windows 3.1  
 SOFTWARE: WORD PERFECT 6.1 and ASCII  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/603,753D  
 FILING DATE: 20 FEB 1996

CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: U.S. 08/373,799  
FILING DATE: 17 JAN 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: ARLES A. TAYLOR, JR.  
REGISTRATION NUMBER: 39,395  
REFERENCE/DOCKET NUMBER: 1242/2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919) 493-8000  
TELEFAX: (919) 419-0383  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3418  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHEICAL: no  
ANTI-SENSE: no  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens sapiens  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE: adult  
TISSUE TYPE: female breast  
CELL TYPE: normal breast tissue  
CELL LINE: HMEC  
ORGANELLE: no  
FEATURE:  
NAME/KEY: BRCA2 protein  
LOCATION: 1 to 3418; Genbank locus HS043746  
IDENTIFICATION METHOD:  
OTHER INFORMATION: BRCA2 protein has a negative  
OTHER INFORMATION: regulatory effect on growth of human mammary cells.  
PUBLICATION INFORMATION:  
AUTHORS: Wooster, R. et al.  
TITLE: Identification of the breast cancer  
JOURNAL: Nature  
VOLUME: 379  
PAGES: 789-792  
DATE: 1995  
RELEVANT RESIDUES IN SEQ ID NO: 4: granin box  
RELEVANT RESIDUES IN SEQ ID NO: domain at amino acids 3334-3344  
US-08-603-753D-4

218 KQ-----KPISNYYVYSESNCEIKDIIRDOQYVNLIGTSDKPYVYTDVLAISNL--- 270  
DB 485 KQALISGTSFVASSFGIIRKSIIFRIRESKP-ETFNASFSGHMTDPNFKKETASBSGLEIH 543  
QY 271 --CDEKQ-KVAVEVIKN-----LITNTLVLDLGLGLTPANKXGIAHLAKSSNFY--- 318  
QY 22 NGNEVKLINLIKD-----VLPTQVSGYNIETEPD---CYSDAISLQSLPDVFSRDS 69  
DB 270 SGNSEFK-VNSCKDHGKMPNLEDEYEVYVDTSEDSFSLCFSKCRKTNLQKVRTISKT 328  
QY 70 ---IFLPYLVISGKYSLDE-----SLVRGV---TGDLS-----FVSSS 103  
DB 329 RKKIF-----HEANDECEKSKNQVKEKYSFVSEVERPNDTDLDSNVVAHQKPFESGS 380  
QY 104 ASVNGSVYGFQYLC--SNFLISSPENGTO-QASSLLELA---QKGYEQIYVPDVASSSS 157  
DB 381 DKISKEV--VPSLACEWQSLTSLGNGAQMEKIPLIHISCDQNISEXDLDTENKRRKD 438  
QY 158 FTFVGLVQQLQSSSAVNDIKASDLPQSGDQVNDITQKYRTILDSITVVASQREYINSV 217  
DB 439 F-----LTSNSLP--RISLSPKSEKPLNEETVANKRD--EQHQLSHSTDCILAV 484  
QY 218 KQ-----KPISNYYVYSESNCEIKDIIRDOQYVNLIGTSDKPYVYTDVLAISNL--- 270  
DB 485 KQALISGTSFVASSFGIIRKSIIFRIRESKP-ETFNASFSGHMTDPNFKKETASBSGLEIH 543  
QY 271 --CDEKQ-KVAVEVIKN-----LITNTLVLDLGLGLTPANKXGIAHLAKSSNFY--- 318

DB 544 TVCSQKEDSLCPNLIIDNGSWPATTONSVALKXNAGLISTLKKTNTKIYAIHDETFYK 603  
QY 319 -----AOLSOQFDAKESEVRVLRVDFPANK 344  
DB 604 KIPKQKSELINCSAQFEANAFAEP-----LTFANAD 635

Search completed: January 24, 2003, 19:43:37  
Job time: 25 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 24, 2003, 19:37:09 ; Search time 38 Seconds

(without alignments)  
1248.348 Million cell updates/sec

Title: US-09-675-509-4

Perfect score: 1802  
Sequence: 1 MSTOPKTLTVGLFPYLPWN.....CVDPAKEVNCAGVLRPL 356

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002:.\*  
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	139.5	7.7	10182	23	ABP38314
2	138.5	7.7	5024	22	AA682935
3	105.5	5.9	511	9	AA61161
4	104.5	5.8	511	9	AA61180
5	104.5	5.8	512	11	AA607574
6	104	5.8	753	23	ABP29819
7	104	5.8	751	23	ABP27624
8	103.5	5.7	397	23	ABP39259
9	103.5	5.7	4643	22	ABP71609
10	103	5.7	286	23	ABP53627

11	101.5	5.6	747	23	ABP90996
12	101.5	5.6	859	22	ABP71182
13	101.5	5.6	1787	23	ABP4791
14	101	5.6	346	15	AA60654
15	100.5	5.6	745	19	AA673024
16	100.5	5.6	745	19	AA673024
17	100	5.5	346	15	AA60651
18	99.5	5.5	856	23	ABP90898
19	99	5.5	889	23	ABP92551
20	98.5	5.5	1149	14	AA632354
21	98.5	5.5	1149	15	AA631692
22	98.5	5.5	1161	23	ABP5312
23	98.5	5.5	1239	20	AA604734
24	98.5	5.5	1373	20	AA604730
25	98.5	5.5	2000	20	AA604732
26	98.5	5.5	2053	22	ABP71118
27	98	5.4	408	21	AA635740
28	98	5.4	456	21	AA635739
29	98	5.4	474	21	AA635738
30	97.5	5.4	624	19	AA668204
31	97.5	5.4	639	23	ABP40816
32	97.5	5.4	2037	21	AA653753
33	97.5	5.4	2070	20	AA604733
34	97.5	5.4	2701	17	AA699850
35	97	5.4	461	23	ABP64816
36	97	5.4	461	23	AA696176
37	97	5.4	487	22	AA681291
38	97	5.4	615	21	AA624068
39	97	5.4	615	21	AA66678
40	97	5.4	615	22	AA665201
41	97	5.4	644	22	AA670158
42	97	5.4	1075	22	AA630804
43	96.5	5.4	470	21	AA679305
44	96.5	5.4	2662	23	ABP91373
45	96	5.3	637	23	ABP49873

#### ALIGNMENTS

RESULT 1  
ABP38314 standard; Protein; 10182 AA.  
XX  
AC ABP38314;  
XX  
DT 24-UTL-2002 (first entry)  
XX  
DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3159.  
XX  
DE Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;  
XX  
KW antibacterial; gene therapy.  
XX  
OS Staphylococcus epidermidis.  
XX  
PN US6380370-B1.  
XX  
PD 30-APR-2002.  
XX  
PF 13-AUG-1998; 98US-0134001.  
XX  
PR 14-AUG-1997; 97US-055779P.  
XX  
PR 08-NOV-1997; 97US-064964P.  
XX  
PI (GENO-) GENOME THERAPEUTICS CORP.  
XX  
PI Doucette-Stamm LA, Bush D;  
XX  
DR WPI; 2002-381255/41.  
XX  
DR N-PSDB; ABP90859.  
XX  
PT Novel isolated nucleic acid encoding a Staphylococcus epidermidis  
polypeptide, useful for diagnosing and treating bacterial infections -

XX		PS	Diclosure; SEQ ID 3159; 267bp; English.
CC	ABN90538	to ABN93374	represent Staphylococcus epidermidis open reading frame (ORF) nucleic acid sequences which encode the amino acid sequences given in ABP5124 to ABP37960. The S. epidermidis sequences have antibacterial activity and can be used in gene therapy. The sequences can also be used in the diagnosis and treatment of bacterial infections, particularly S. epidermidis infections. The sequences can be used to screen for compounds able to interfere with the S. epidermidis life cycle or inhibit S. epidermidis infection.
CC	N.B.	The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the USPTO web site.	
XX	Seq	Sequence	10182 AA;
Query Match:	Best Local Similarity	7.7%; Score 139.5; DB 23; Length 10182.	
ches	89; Conservative	59; Mismatches 140; Indels 147; Gaps 17,	
OY	24	NEVKLINIKDVLPROVSGYNIEYEFCISDASLQSLPDVSTSDSIFLYLVSLGVKS	83
DB	8873	NETQAINNARDAL-NKTQGQNLDFNAIDPFKD-----DIKTXD-----ALNGIER	8917
OY	84	L-----DESLVRGVTGDLSHFVSSASVNGSV-----	110
DB	8918	LTPAAKSKAEKLIDSLKEFNKAQFTHANDIMNTNSIAQLSRIVNOAFPUNDAMKSLRDEL	8977
OY	111	--YGFPQYLCSNFLSSPNGTOQASSLLLAOKV-----GYEOLVYPDVA	153
DB	8978	NNQAFPVQASNVYINSEDLKKQFDHALNARKVLAKENGNLDEKQIQGLKQVI-----	9032
OY	154	SSSSSFTEFLYQ-----QLQS-----SSSAVDIKASD	182
DB	9033	EDTKDALNGIQRLSKAKAKAIQYQGISLYINDAQRHIAENNHNINSDDLSSLANTLSKAD	9092
OY	183	LPOSGOVNQKIDITQKRRTLDSTFWVASQREYINSVGKRPISYYVYGSESWEIMDIR	242
DB	9093	L---DNAMKDLDTIES--NSTSVSNVYINADK-----NLQLEPDEAL-----	9132
OY	243	DOOVNYQLIGTSPDKPYVTVTLAINSNLCDEK-----QKAVEVINKULTLTNTLVLDLG	296
DB	9133	-QQASATTSKTSSENPAETIEVGLSQAIYDTKNAALNGEORLATEXSKDCLKIKGLMDLK	9191
OY	297	LGLTLPANK---NGIAHLAKSSNFYAQLSQGFDAKESEVRVL-----RCVDF--	340
DB	9192	AQLEDVTNKVNSANTLETLSQLTGSTLELNDKKKLRLDKLTLPVPKASLANRYNRADYNL	9251
OY	341	---ANKEVNKCAGVL	352
DB	9252	KROFNKALKEAKGVL	9266
Result 2			
ID	AAG82935	standard; Protein; 5024 AA.	
AC	AAG82935;		
DT	03-SEP-2001	(first entry)	
DE	S. epidermidis	open reading frame protein sequence SEQ ID NO:2964.	
KW	Staphylococcus epidermidis	SRI strain; infection; diagnosis;	
KM	vaccination; endocarditis.		
OS	Staphylococcus epidermidis.		
PN	WO200134809-A2.		
PD	17-MAY-2001.		

Query Match	7.7%	Score 138.5	DB 22	Length 5024
Beet Local Similarity	20.5%	Pred. No. 0.075		
Matches 89	Conservative 59	Mismatches 140	Indels 147	Gaps 17
XX 09-NOV-2000	2000MO-US30782			
XX 09-NOV-1999	99US-0164258			
XX (GLAX )	GLAXO GROUP LTD			
XX Kimerly MJ				
XX WPI, 2001-316495/33				
XX N-PSDB, AAH53785				
XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,				
XX useful for vaccinating against infections, e.g. endocarditis -				
XX Claim 18, Page 779-781, 2186pp; English.				
XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides				
XX (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.				
XX (I) and (II) can have antibacterial activity and therefore can be used				
XX in vaccination. The nucleic acid (I) may be used to produce the				
XX S. epidermidis polypeptides (II) via the production of vectors				
XX containing them which are used to produce host cells which express the				
XX polypeptides. The polypeptides (II) (and/or nucleic acids) may then be				
XX used to vaccinate subjects and to raise antibodies against the bacteria.				
XX The polypeptides may also be used to assay for other inhibitors of their				
XX activity and therefore identify compounds that may be used for the				
XX treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to				
XX AAH55090 represent specifically claimed S. epidermidis genomic DNA				
XX polynucleotide sequences from the present invention. AAH55091 to				
XX AAH55098 represent oligonucleotide sequences and primers which are used				
XX in the exemplification of the present invention.				
XX N.B. The present invention specifically claims all the polynucleotide				
XX sequences given in the sequence listing of the present specification,				
XX however the sequence listing only goes up to SEQ ID NO:4454, so even				
XX though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,				
XX no sequences are present for SEQ ID NO:4455 to 4464.				
XX Sequence 5024 AA;				
QY 24 NEVKILNLIKVLPRQVSGVNIETREPFQCYASLSQSLPQVSTQSIPLPYLSLGVKS 83				
DB 3927 NFTAQINNARDAL-NKTOGQNDLDFNAIDTFKQ-----DIFKTD-----ALNGIER 3971				
QY 84 L-----DESLRGVGTGDIHSFVSSASVNGSV----- 110				
DB 3972 LTNAAKSKAEKLDLSKFTNKQAFTHANDEIMNTNSIAQSLRYVNGAFDNDAMKSLRDEL 4031				
QY 111 --YGFPPQYLCNLFSSPNGTQQAASSLLEAQRV-----GYEQIYVPDVA 153				
DB 4032 NNQAFPVQAASSVYINSDDELKQOFPHALSNARKVLAKEGKNLDEIQIEGLKQVI----- 4086				
QY 154 SSSSFVPELYQ-----QLLOS-----SSSAVDIKASD 182				
DB 4087 EDTKDALNGIQFLSKAKAKAIGVYOSLSYINDAQRHIAESNTHNSDDLSLANTLSKASD 4146				
QY 183 LPQSGQVQNKDITQKKRTILDTVVAASQREYINSVKQGRPISNYVGVGESMCEIYDIR 242				
DB 4147 L----DNAMKDLRDLIES--NSTSPVNSVYINADK-----NLQIEPFEAL----- 4186				
QY 243 DOQYNVOLIGTSDKPYVYTDVLALNSNLCDK-----QKVAVEVIKNLLTNTVLVDLLG 296				
DB 4187 -QQASATSSKTSSENPAITIEVIGLSQAIVDTKNALNGEQRALATEKSKDKLTKGLKDUNK 4245				
QY 297 LGITLTPANK---NGIAHLAKSNFPAQSLQGFDAKESVRVY-----RCDFE-- 340				
DB 4246 AQLIEDVTIVNSANTLTETLSQLTQSTLKLNDKMLLRDXLKTLLVNPVKASLNYRNDVNL 4305				
QY 341 ---ANKEVKNCAQVLT 352				

Db 4306 KROPNKALKEAKGV 4320

## RESULT 3

ID AAP81161 standard; protein; 511 AA.

AC AAP81161;

DT 12-OCT-1990 (first entry)

DE Recombinant alpha-amylase.

KW Alpha amylase; yeast; brewing; baking.

OS Schwanniomycetes castellii ATCC 26076.

PN EP257115-A.

PD 02-MAR-1988.

PF 21-AUG-1986; 86EP-0111586.

PR 21-AUG-1986; 86EP-0111586.

PA (HETB) HEINEKEN TECHNISCH BEHEER BV.

PI Straesser A, Martens FB, Dohmen J, Hollenberg CP;

DR WPI: 1988-057532/09.

DR N-PSDB; AAN81477.

PT Prod. of amylolytic enzymes by transformed microorganisms -  
 PT comprises use of recombinant DNA technology to give transformant  
 PT having DNA from donor yeast.

PS Disclosure; P; English.

CC The sequence encoding the protein can be used to construct an  
 CC expression vector for the prodn. of alpha amylase in yeast cells.  
 CC The transformants and recombinant enzyme produced can be used in  
 CC fermentation processes e.g in baking and brewing. The transformed  
 CC yeasts are esp. suitable for prodn. of low carbohydrate beers.

SQ Sequence 511 AA;

Query Match 5.9%; Score 105.5; DB 9; Length 511;

Best Local Similarity 22.0%; Pred. No. 1.9;  
 Mismatches 80; Conservative 48; Indels 117; Gaps 21;

```

QY 18 SMNENGNEVKLILKDVLPQTQ-----VSGYNIETFEPCYSDASLOSLEDPVST 67
Db 159 AMNGDSSVD---YSSFTFPNOQSYFHDYCLITNYNDQTNVEDCWEQDTEVSLPDLSTE 214
QY 68 DS-----IFLPYLVSLGVSGLD-----ESLVRGVTGDLHSFVSSASVN-GSVY-GFPQYL 117
Db 215 DNEVIGVFTWVSDPVQVNSIDGLRIDSAGHVDTSLTKEFEDASGVNNGEYVQGDPTT 274
QY 118 C-----SNFLSSP-----NGTQASSLLELAQV-----GYE 145
Db 275 CPYQNYMKGVNTNPLYVPYRPFSDTSATSS--ELTSMISTLQSSCDVSLGNFIENHD 332
QY 146 QIVPDPVASSSS-----FTVFG-----LY-----QQLQSSSSA-----VDIK 179
Db 333 QVRPFVTSVSDTSLIKDMAFIIIGDGIPTIYGGQGLNGSGDPANREALMLSGYNTDSE 392
QY 180 ASDLPQSDQVNDITOK---YRTILDSTVVASQREYINSVK-----218
Db 393 YVELIKLQNIQIRQAIKDSAVST-YKSSVSSSDHYIAIRKSDANQLISIFNNLSNG 451
QY 219 -GKPISSNYVYGSESMCEIKIIR-----DQYNNQLISTSKPYYTVDVALNSN 269
Db 452 GSDITVSN--TGYS-SGDKVIDIISCNVLAGDSGLSVSISGRMPQVYAPSSVL-SGS 507

```

QY 270 LCDE 273

Db 508 ICNQ 511

## RESULT 4

ID AAP81180 standard; protein; 511 AA.

AC AAP81180;

DT 12-DEC-1990 (first entry)

DE Sequence of alpha-amylase encoded by genomic DNA of Schwanniomycetes  
 DE castellii.

KW Brewing; beer; breadmaking; bioass.

OS Schwanniomycetes castellii.

PN EP260404-A.

PD 23-MAR-1988.

PF 17-JUL-1987; 87EP-0110370.

PR 17-JUL-1987; 87EP-0110370.

PA (HETB) HEINEKEN TEC BEHEER NV.

PI Straesser A, Martens FB, Dohmen J, Hollenberg CP;

DR WPI: 1988-078794/12.

DR N-PSDB; AAN81525.

PT Polypeptide prodn. from cells transformed with yeast DNA -  
 PT esp. coding alpha amylase or glucoamylase, able to convert  
 PT starch to ethanol

PS Claim 18; Fig 2A-2C; 74pp; English.

CC A cosmid library is constructed from genomic DNA of ATCC26076 and this  
 CC screened for inserts contg. the Aa gene by ability to transform  
 CC S.cerevisiae GRF 18 to histidine prototrophy. One positive cosmid  
 CC (32kb; pyc1) was digested with EcoRI, religated and used to transform  
 CC E.coli JAZ21. Positive transformants contain pyc1-alpha, which includes  
 CC a 5kb EcoRI fragment contg. the gene in AAN81525.

SQ Sequence 511 AA;

Query Match 5.8%; Score 104.5; DB 9; Length 511;

Best Local Similarity 22.7%; Pred. No. 2.3;  
 Matches 83; Conservative 44; Mismatches 116; Indels 123; Gaps 22;

```

QY 18 SMNENGNEVKLILKDVLPQTQ-----VSGYNIETFEPCYSDASLOSLEDPVST 67
Db 159 AMNGDSSVD---YSSFTFPNOQSYFHDYCLITNYNDQTNVEDCWEQDTEVSLPDLSTE 214
QY 68 DS-----IFLPYLVSLGVSGLD-----ESLVRGVTGDLHSFVSSASVN-GSVY-GFPQYL 117
Db 215 DNEVIGVFTWVSDPVQVNSIDGLRIDSAGHVDTSLTKEFEDASGVNNGEYVQGDPTT 274
QY 118 C-----SNFLSSP-----NGTQASSLLELAQV-----GYE 145
Db 275 CPYQNYMKGVNTNPLYVPYRPFSDTSATSS--ELTSMISTLQSSCDVSLGNFIENHD 332
QY 146 QIVPDPVASSSS-----FTVFG-----LY-----QQLQSSSSA-----VDIK 179
Db 333 QVRPFVTSVSDTSLIKDMAFIIIGDGIPTIYGGQGLNGSGDPANREALMLSGYNTDSE 392
QY 180 ASDLPQSDQVNDITOK---YRTILDSTVVASQREYINSVK-----218
Db 393 YVELIKLQNIQIRQAIKDSAVST-YKSSVSSSDHYIAIRKSDANQLISIFNNLSNG 451
QY 219 -GKPISSNYVYGSESMCEIKIIR-----DQYNNQLISTSKPYYTVDVALNSN 269
Db 452 GSDITVSN--TGYS-SGDKVIDIISCNVLAGDSGLSVSISGRMPQVYAPSSVL-SGS 507

```

```

Db      393 YVELISKLNQIRNOAIKDSAYST-YKSSVSSSDHYIAIRKSDANQIISFNNGSNG 451
Oy      219 -QGKRISNYVYGSEBSMCEIKDIIIPDOYVNLIGTSDK-----PYVYTDVIALN- 267
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      432 SODITVSN--IGYS-SGDKVIDII---SCNSVLGSDGSLSVISGMPQVYAPSSVLSA 505
Oy      268 SNLCDE 273
          ||| |||
Db      506 SGICNQ 511

RESULT 5
AA07574
ID      AA07574 standard; protein, 512 AA.
AC      AAR07574;
DT      02-FEB-1991 (first entry)
XX      D
XX      Alpha-amylase encoded by the AMY1 gene.
XX      Schwanmioncyes yeast cells; expression cassette; alpha-amylase;
KW      marker gene; region; signal peptide; terminator;
KW      autonomously replicating sequence.
XX      Schwanmioncyes occidentalis.
OS
FH      Key location/Qualifiers
FH      Disulfide-bond 66..74 /note="1"
FT      Disulfide-bond 186..200 /note="2"
FT      Disulfide-bond 276..319 /note="3"
FT      Disulfide-bond 475..510 /note="4"
FT      EP394538-A.
XX      31-OCT-1990.
XX      PD
XX      PF 28-APR-1989; 89EP-0107780.
XX      PR 28-APR-1989; 89EP-0107780.
XX      PA (RHEI ) RHEIN BIOTECH GBS.
XX      PF Hollenberg C, Strasser A;
XX      MPI; 1990-328670/44.
XX      DR N-PSDB; AAQ06388.
PT      Transformed Schwanmioncyes yeast cells - contg. an expression
PT      cassette contg. region, DNA coding for foreign protein and
PT      terminator
XX
XX      Disclosure; Fig 1B(1-6); 59pp; English.
XX
CC      The expression cassette may contain a region (R), a signal peptide
CC      sequence (S), a foreign protein sequence (F) and a terminator (T)
CC      of which R and/or S and/or T are derived from the alpha-amylase gene
CC      of S. occidentalis.
CC      R is pref. a 1.8 kb BglII-XhoI fragment (bases -1 to -540 of this
CC      sequence).
CC      S contains all or part of one of the sequences represented
CC      in AAQ06383-87.
CC      F is e.g. a cellulase, interleukin, insulin-like-growth factor,
CC      interferon etc.
CC      T is pref. all or part of the terminator of this sequence
CC      (bp 1537-1740).
CC      AMY1 acts as selective marker for the vector carrying the expression
CC      cassette.
CC      See also AAQ06389.
CE

```

[illegible]

PT	RESULT 6
XX	ABP29819
XX	ID ABP29819 standard; Protein; 753 AA.
XX	AC ABP29819;
XX	DT 02-JUL-2002 (first entry)
DE	Streptococcus polypeptide SEQ ID NO 8814.
XX	Streplococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KM	Group A streptococcus; Streptococcus pyogenes; antibacterial;
KW	antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX	Streptococcus agalactiae.
OS	
XX	WO200234771-A2.
PN	
XX	
PD	02-MAY-2002.
XX	
Pf	29-OCT-2001; 2001WO-GB04789.
XX	
PR	27-OCT-2000; 2000GB-0026333.
PR	24-NOV-2000; 2000GB-0028727.
PR	07-MAR-2001; 2001GB-0005640.
XX	
PA	(CHIR-) CHIRON SPA.
PA	(GENO-) INST GENOMIC RES.
XX	
PI	Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
PI	Tetreltin H;
XX	
DR	WPI; 2002-352536/38.
DR	N-FSDB; ABN70450.
XX	
PT	New Streptococcus protein for the treatment or prevention of infection or disease caused by Streptococcus bacteria, such as meningitis, and

for detecting a compound that binds to the protein -

XX PS Claim 1; Page 3998; 4525pp; English.

XX CC The invention relates to a protein (ABP25413-ABP30895) from group B

XX CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS

XX CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in

XX CC the specification. The proteins have antibacterial and antiinflammatory

XX CC activity. (1), nucleic acids encoding (1), ABN66044-ABN71526 and

XX CC antibodies that bind (1) are used in the manufacture of medicaments for

XX CC the treatment or prevention of infection or disease caused by

XX CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

XX CC Nucleic acids encoding (1) are used to detect Streptococcus in a

XX CC biological sample. (1) is used to determine whether a compound binds to

XX CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be

XX CC used as a vaccine or diagnostic composition. The disease caused by

XX CC Streptococcus that is prevented or treated may be meningitis. Nucleic

XX CC acid encoding (1) may be used to recombinantly produce (1) and may be

XX CC used in gene therapy. Antibodies to (1) are used for affinity

XX CC chromatography, immunoassays, and distinguishing/identifying

XX CC Streptococcus proteins.

XX SQ Sequence 753 AA;

Query Match 5.8%; Score 104; DB 23; Length 753;

Best Local Similarity 23.0%; Pred. No. 4.4;

Matches 59; Conservative 46; Mismatches 76; Indels 76; Gaps 12;

QY 111 YGFPQ---YLCNPLSSPNGTQOASLLLAQKVGEGVYPPDVASS-SFTVFGLYQ 165

DB 319 YGYKSSKITYSAHYALYS---EIDAKELQKXIAKDNVNEIKSMNLSPIFNQYL 374

QY 166 QLLQSSSAVDIKAS---DLPOSGD-----QVNKDITOKRTILDS 204

DB 375 QLLQIVISSEINLKALDNTVDLPIENNFTLDIQYNKLDTAIKSLKFTKYQOEVRKA 434

QY 205 TVVASOREYINSVKQKRPISNYVGYSESMECEIKDIRDOQYVNLIGTSDK----- 256

DB 435 TKYSKKELVNA-ELTVKISN-----DNILLDMQAISNNGSTKPFYSVKRL 481

QY 257 PYVYTDVIALNSN-LCEKQKVAVEIKNLTNTLVLDLGLGLTPANKNGIAHLAKSS 315

DB 482 PYVPQVIMTTTSNVLMPQKQ---VEKVKLTLPPFI-----SNKEVLNHDLSIVE 526

QY 316 NFVAQLSQQPDAKESV 332

DB 527 N-----DAOKOKV 534

RESULT 7

ABP27624

ID ABP27624 standard; Protein; 761 AA.

XX AC ABP27624;

XX DT 02-JUL-2002 (first entry)

XX DE Streptococcus polypeptide SEQ ID NO 4424.

XX KM Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;

XX KM group A streptococcus; Streptococcus pyogenes; antibacterial;

XX KM antiinflammatory; infection; vaccine; meningitis; gene therapy.

XX OS Streptococcus agalactiae.

XX PN WO200234771-A2.

XX PD 02-MAY-2002.

XX PF 29-OCT-2001; 2001MO-GB04789.

XX PR 27-OCT-2000; 2000GB-0026333.

XX PR 24-NOV-2000; 2000GB-0028727.

PR 07-MAR-2001; 2001GB-0005640.

XX PA (CHIR-) CHIRON SPA.

XX PA (GENO-) INST GENOMIC RES.

XX PI Telford J, Maignani V, Margart Roe YI, Grandi G, Fraser C;

XX PI Tettein H;

XX DR WPI; 2002-352536/38.

XX DR N-PSDB; ABN68255.

XX PT New Streptococcus protein for the treatment or prevention of infection

XX PT or disease caused by Streptococcus bacteria, such as meningitis, and

XX PT for detecting a compound that binds to the protein -

XX PS Claim 1; Page 3603; 4525pp; English.

XX CC The invention relates to a protein (ABP25413-ABP30895) from group B

XX CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS

XX CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in

XX CC the specification. The proteins have antibacterial and antiinflammatory

XX CC activity. (1), nucleic acids encoding (1), ABN66044-ABN71526 and

XX CC antibodies that bind (1) are used in the manufacture of medicaments for

XX CC the treatment or prevention of infection or disease caused by

XX CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

XX CC Nucleic acids encoding (1) are used to detect Streptococcus in a

XX CC biological sample. (1) is used to determine whether a compound binds to

XX CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be

XX CC used as a vaccine or diagnostic composition. The disease caused by

XX CC Streptococcus that is prevented or treated may be meningitis. Nucleic

XX CC acid encoding (1) may be used to recombinantly produce (1) and may be

XX CC used in gene therapy. Antibodies to (1) are used for affinity

XX CC chromatography, immunoassays, and distinguishing/identifying

XX CC Streptococcus proteins.

XX SQ Sequence 761 AA;

Query Match 5.8%; Score 104; DB 23; Length 761;

Best Local Similarity 23.0%; Pred. No. 4.5;

Matches 59; Conservative 46; Mismatches 76; Indels 76; Gaps 12;

QY 111 YGFPQ---YLCNPLSSPNGTQOASLLLAQKVGEGVYPPDVASS-SFTVFGLYQ 165

DB 327 YGYKSSKITYSAHYALYS---EIDAKELQKXIAKDNVNEIKSMNLSPIFNQYL 382

QY 166 QLLQSSSAVDIKAS---DLPOSGD-----QVNKDITOKRTILDS 204

DB 383 QLLQIVISSEINLKALDNTVDLPIENNFTLDIQYNKLDTAIKSLKFTKYQOEVRKA 442

QY 205 TVVASOREYINSVKQKRPISNYVGYSESMECEIKDIRDOQYVNLIGTSDK----- 256

DB 443 TKYSKKELVNA-ELTVKISN-----DNILLDMQAISNNGSTKPFYSVKRL 489

QY 257 PYVYTDVIALNSN-LCEKQKVAVEIKNLTNTLVLDLGLGLTPANKNGIAHLAKSS 315

DB 490 PYVPQVIMTTTSNVLMPQKQ---VEKVKLTLPPFI-----SNKEVLNHDLSIVE 534

QY 316 NFVAQLSQQPDAKESV 332

DB 535 N-----DAOKOKV 542

RESULT 8

ABP39259

ID ABP39259 standard; Protein; 397 AA.

XX AC ABP39259;

XX DT 24-JUL-2002 (first entry)

XX DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:4104.

XX KM Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;

KM antibacterial; gene therapy.  
XX  
OS Staphylococcus epidermidis.  
XX  
PN US6380370-B1.  
XX  
PD 30-APR-2002.  
XX  
PF 13-AUG-1998; 98US-0134001.  
XX  
PR 14-AUG-1997; 97US-055779P.  
PR 08-NOV-1997; 97US-064964P.  
XX  
PA (GENO-) GENOME THERAPEUTICS CORP.  
XX  
PI Doucette-Stamm LA, Bush D;  
XX  
DR WPI: 2002-381255/41.  
DR N-PSDB; ABN91804.  
XX  
PS Novel isolated nucleic acid encoding a staphylococcus epidermidis polypeptide, useful for diagnosing and treating bacterial infections -  
XX  
XX Disclosure; SEQ ID 4104; 267pp; English.  
XX  
CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading  
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences  
CC given in ABP5124 to ABP37960. The S. epidermidis sequences have  
CC antibacterial activity and can be used in gene therapy. The sequences  
CC can also be used in the diagnosis and treatment of bacterial infections,  
CC particularly S. epidermidis infections. The sequences can be used to  
CC screen for compounds able to interfere with the S. epidermidis life  
CC cycle or inhibit S. epidermidis infection.  
CC N.B. The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from the  
CC USPTO web site.  
XX  
SQ Sequence 397 AA;  
XX  
Query Match 5.7%; Score 103.5; DB 23; Length 397;  
Best Local Similarity 18.5%; Pred. No. 1.9;  
Matches 74; Conservative 72; Mismatches 142; Indels 111; Gaps 18;  
XX  
QY 52 CYSASLQSLDVPF---STDSIFLPYLVSLG-VKSLDE-----SLVRGV 92  
DB 1 CYNDLITENNEVEFYAMKSTFFFMNIYAMGGTVKSVTLANTLAEKGPVTIISVFRGA 60  
QY 93 TG---DLHSFV-----SSSASVNGSVYGFPOYLCNLFLLSSPNCQOASSLL 136  
DB 61 DSPYFELHSAIKKVVVDYRLKLNKTRAITANRIKKYTPFLNTKVISOFEPEKSOFSYV 120  
QY 137 ELAOKVGEQIVYPD---VASSSFTVFGLYQQLQSSSA---AVDIKASDLPOSGDQ 189  
DB 121 E---KKMKAIKRIHTKTDVLCGRASFNIT-----LISKAKAEIYTIMMEHNPAHBDQ 171  
QY 190 VNKDITQKRTI--LDSTVVASQREY-----INSVKQGPIS----- 224  
DB 172 YQKEIIAAYRNINIKITTLTVADQKYOQSLKTPYVVIPIPMVTEKRIAPKPKRIISAGRL 231  
QY 225 NYVYGV---ESMCEIKIIRDOQYNNVOLIGTSDKP-----YVTTDVLA 265  
DB 232 EYKGVYDILLLESIRLIQDLRLQNLNDVHIYSGSKTSLVDFINQYHNDLIKIEPTQE 291  
QY 266 LNSNLCDEK-----QKVAVEVIKNLLTNTLVLDLG-LGLTLPAKNGIAHLAKSSN 316  
DB 292 LNNKLAQKIVVPSRNQEGFMIIIEAMVQDVIIVISFEGNVPDSITINNNGNGYLVNTEN 351  
QY 317 FYAQLSQGFDAKESSEVRVLCVDFPANKVEKNCAGVLRPF 355  
DB 352 -VSELAKRIDLTTQHYNELDHI-----IENSKDTLKQF 383  
RESULT 9

ABB71609  
ID ABB71609 standard; Protein; 4643 AA.  
XX  
XX ABB71609;  
AC  
XX  
XX 26-MAR-2002 (first entry)  
DT  
XX  
XX Drosophila melanogaster polypeptide SEQ ID NO 41619.  
DE  
XX  
XX Drosophila; developmental biology; cell signalling; insecticide;  
KM pharmaceutical.  
KM  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
PD  
XX  
PF 23-MAR-2001; 2001WO-US09231.  
XX  
XX 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX  
XX (PEKE ) PE CORP NY.  
PA  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
XX WPI: 2001-656860/75.  
DR N-PSDB; ABL15712.  
DR  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
XX Disclosure; SEQ ID NO 41619; 21pp + Sequence Listing; English.  
XX  
PS  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA  
CC sequences (AB101840-AB116175) and the encoded proteins  
CC (ABB57737-ABB72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 4643 AA;  
XX  
Query Match 5.7%; Score 103.5; DB 22; Length 4643;  
Best Local Similarity 20.9%; Pred. No. 70;  
Matches 78; Conservative 50; Mismatches 110; Indels 135; Gaps 19;  
XX  
QY 2 STQKTLTVGLFPYLPKMNENGENEVLINLIKDVLPYOVSGYN-----IETTEP-- 51  
DB 3531 NTTYTFEDI-----RSGNEGGLFRLQDQSLRTASRFNNLQDEFYIOVRVFDNG 3580  
QY 52 ---CYSASLQSLDVPFSTDSIFLPYLVSLGVSLSLSESLRGVYTGDLHSVSS-----S 103  
DB 3581 TPPLYSDA-----WVV---VXIIESQYPPVTPPLEVTINSFEDDFS 3619  
QY 104 ASVNGSVYGFPO--YLCNLF--LLSSPNCQOASSLLELAOKVGEQIVYPD---DV----- 152  
DB 3620 GAFIGKHAASQDXYDELNFEVLVSGPDDMYSSKLFNINSNTGK---IYALSNDIGLYK 3676  
QY 153 ---ASSSFTVFGLYQQLQSSSAADVIRKASDLPOSGDVNKDITQKRTIISDVVA 208  
DB 3677 LNVGSQCKFHVFSI-----VKINVELVTNMLKESVYIRFRIRASSEFL 3722  
QY 209 SQRE-YINSVQGKRPISNYVYGVSESMCEIKDIR-----DQ 244  
DB 3723 SHRTFMRSIR-----NIMRCROKDVILITLQSDYOKASOHAVGNRRARSIDS 3770

OY 245 QYNV-----OLIGTSDKRYVTDTVLATNSNLCDEKQVAEVIKNLLTNLTLYVDLGL 297  
| | | | | : : : : :  
Db 3771 DLANVFAVRKQIILPDSDEFSTDEL---RQLTLIDKGNIEENE-----TNLVEDV--L 3819  
| | | | | : : : : :  
OY 298 GLTLPANKNGIAH 310  
| | | | | : : : : :  
Db 3820 PSTGSKNKNDGVH 3832  
| | | | | : : : : :

RESULT 10  
ABB53627  
ID ABB53627 standard; Protein; 286 AA

DT 16-MAY-2002 (first entry)  
XX  
DE Lactococcus lactis protein plpD.  
XX  
XX biosynthesis; biodegradation; lactic bacterium; yogurt; cheese  
XX  
XX Lactococcus lactis IL1403.

PF 11-APR-2000; 2000PR-0004630.  
XX  
PR 11-APR-2000; 2000PR-0004630.  
XX  
PA (INRG ) INRA INST NAT RECH AGRONOMIQUE

PI	Bolotline A,	Sorokine A,	Renault P,	Ehrlich SD,
XX				
DR	WPI; 2002-043418/06.			

PT New nucleotide sequence useful in the identification or *Lactococcus*  
PT *lactis* and related species -

PS Claim 6; SEQ ID No 329; 2504pp; French

CC The present invention is related to a Lactococcus lactis nucleotide  
CC sequence (AB390521) and related proteins (AB553300-AB55621). The  
CC nucleic acid sequence is useful in the detection of and/or amplification of  
CC nucleic acid sequence, particularly to identify Lactococcus lactis or  
CC related species. The proteins of the invention are useful for the  
CC biosynthesis or biodegradation of a composition of interest. The  
CC invention helps research in lactic bacteria, particularly useful in the  
CC production of yogurt and cheese.  
CC Note: The sequence data for this patent is based on equivalent patent  
CC WO2001/77334 (published 18-Oct-2001) which is available in electronic  
CC format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).  
XX  
XQ Sequence 286 AA;

Query Match	5.7%	Score 103;	DB 23;	length 286;
Best Local Similarity	20.7%	Pred NO. 1.3;		
Matches 44;	Conservative 41;	Mismatches 88;	Indels 40;	Gaps 8

QY	13	PFPYLSNNE--NGSEVKLINLIKVLTPVGSNGVNIEXTPEFCSDASLOSIPDVFSDSI	70
Db	99	YAVLSMNNKANGNIVS----IGDTIITPHLYSTKTKANGTEIDKSTIAIPNDVNESR	154
QY	71	FLPLVLVSLGCVKSLDESLVRCGVTDGLHSFVSSASVNGSVYGPQYLCNFFLSSPNGTQ	130
Db	155	AL-YVLKNAGLIKIDTTS--RGALATVMDIRNP-----KGLIITEIDASQ	196
QY	131	QASSLLEIAQRY-GYEQIVPEDVASSSFTVEGFLYQOOLSSSSAAVDIKASDLPQSGDQ	183
Db	197	TPRALDSVAAVIVNNFAISAKIEKES----LYQEPLENDSSQWNIPLAANDSDKNNK	251

```

Qy 190 VAKDITQKYRTILDSTVASQREYINSVKQKP 222
      |:::|      |:::|
Db 252 VYKEVVKAY-----EQKNIADSIKKEYP 274

```

RESULT 11  
ABB90996  
ID ABB90996 standard; Protein; 747 AA

DT	31-MAY-2002	(first entry)
XX		
DE	Herbicidally active polypeptide SEQ ID NO 207	

KW Herbicidal; plant; agriculture; herbicide.

OS *Arabidopsis thaliana*

PN WO200210210-A2

PD 07-FEB-2002

PF 28-AUG-2001; 2001WO-EP09892.

PR 28-AUG-2001; 2001WO-EP09892.

PA (FARB ) BAYER AG

PI Tietjen K, Weidler M,

DR WPI; 2002-269010/31.

PT Identifying plant target proteins for herbicidally active compounds,  
PT comprising aligning and comparing nucleic acid or amino acid sequences  
PT from plant with nucleic acid or amino acid sequences from non-plant  
PT organisms -

PS Claim 5; SEQ ID NO 207; 261pp + Sequence Listing; English.

CC The invention relates to identifying target proteins  
CC (AB896790-AB894016) for hericidically active compounds, comprising  
CC aligning and comparing nucleic acid or amino acid sequences from plant  
CC with nucleic acid or amino acid sequences from non-plant organisms using  
CC suitable search parameters, where plant sequences having an E-value  
CC greater by a factor of 3 than the E-value of most similar non-plant  
CC sequences are selected. The polypeptides or nucleic acids encoding them  
CC are useful for identifying modulators. The identified modulators are  
CC useful as herbicides.

**SQ Sequence 747 AA;**

Query Match	5.6%	Score	101.5	DB	23	Length	747
Best Local Similarity	25.9%	Pred.	No.7.2				
Matches	85	Conservative	48	Mismatches	110	Indels	85
						Gaps	19

DQ  
3 TQKRTLTIVGLFPY-LPSMNNEN---GNEVKLNLTKDVL-----PQVSSVINLEYTEDFPCYSDD 55  
: : | | : | : : : | | : | : | :  
7 SQAFSLTFLFLFELPVSSEQLISSESRTLDELEIKHKLQYPPTLRSMNN--WTFN-CYLP 63

64 SPSFKILCFNGHYTELTVTGNRTVKLPGRFSSDSLFTVLTKLSNKLTL--SLVSLGIS 120

[illegible]

\_\_\_\_\_  
 \_\_\_\_\_  
 ..  
 \_\_\_\_\_  
 ..  
 \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_  
 ..  
 \_\_\_\_\_  
 \_\_\_\_\_  
 ..  
 \_\_\_\_\_  
 ..  
 ..

[illegible]

QY 207 VASQREYINSVKQKPISNYYVG---YSESMCEIKDIRDQYNVQLIGTSDKPYVYTD 262

Db 217 IPEIRKIKLNLQSLDLSSENKFTGSIPRFLSLPSLQNLISLAQNTL--LSGSLP----- 266  
QY 263 VLALNSNLCDKOKAVAEVINKLTLNTL 290  
Db 267 ---NSSLCSKSLKRI-LDVSRLNLTKL 289

RESULT 12  
ABR71182  
ID ABR71182 standard; Protein: 859 AA.  
AC ABR71182;  
DT 26-MAR-2002 (first entry)  
DE Drosophila melanogaster polypeptide SEQ ID NO 40338.  
KM Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.  
OS Drosophila melanogaster.  
XX  
XX WO200171042-A2.  
XX  
XX 27-SEP-2001.  
XX  
XX 23-MAR-2001; 2001WO-US09231.  
XX  
XX 23-MAR-2000; 2000US-191637P.  
XX 11-JUL-2000; 2000US-0614150.  
XX  
XX (PEKE ) PE CORP NY.  
XX  
XX Venter JC, Adams M, Li FWD, Myers EW;  
PI WPI; 2001-656860/75.  
XX  
XX N-PSDB; ABL15285.  
XX  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
XX  
XX Disclosure; SEQ ID NO 40338; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
CC sequences (AB16176-AB130511) and the encoded proteins  
CC (AB16176-AB130511).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX  
XX Sequence 859 AA;  
XX

Query Match 5.6%; Score 101.5; DB 22; Length 859;  
Best Local Similarity 22.5%; Pred. No. 8.9;  
Matches 77; Conservative 46; Mismatches 150; Indels 69; Gaps 14;

QY 43 YNIEKTEFDC-YSDASLQSLPDPVFTSDSIFLPYLVSLGVSGLDLSLVRGVIGLHSPVS 101  
Db 256 YNNETVGTATGEQDSLQOQSPFLVRS--LRFSTSTPSPGSGKIAKR--GKHHPVPV 310  
QY 102 SSASVNGSVYGFPPYLCNPNFLSPNGTQOASLLELAQKVGVEQIVYVDVASSSFTVF 161  
Db 311 ASWMSQEQWAGEPEVMC-----TLQHSINQAEVYK-NY-----TITTSVNC 350  
QY 162 GLYQQLLOSSSAVDAVIRASDLPQGGDVNK-----DITQKRYTILSDSTVVASQR-EYI 214

Db 351 KLVFQLOQOALSLQVHERSERVUSGLQASSLPALAGATQLSHDDPTATLERGVVF 410  
QY 215 NSVQKQPRISNYVYVYSGSMCEIKDIIIRDQYNVOLIGTSDKPYVTVIALNSLQDEK 274  
Db 411 NDAKIERE-----RYEQHEQIRTVSKDTRYSLE-----RQHYINLESLLDVLQKRN 459  
QY 275 OKVAVEVIKNLITNTLVLDL--LGLGLTPANKNGIHLAK-SSNFPVQSLQCF----- 325  
Db 460 TLITRLIFERLVRLVVISISQSCDILLRANIMVATLNMIDYDGFASISDAFVQNEAV 519  
QY 326 -----DAKSEFVRVLR-----CVDPAKNEVKNAGV 351  
Db 520 RTLVLVVDHKKQSSVRALALALATLCCAPQAINQLSCGSI 561

RESULT 13  
ABR49791  
ID ABR49791 standard; Protein: 1787 AA.  
AC ABR49791;  
DT 05-FEB-2002 (first entry)  
DE Listeria monocytogenes protein #2495.  
XX  
XX  
XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;  
KM vitamin B12; bacterial infection; disease.  
XX  
XX Listeria monocytogenes.  
XX  
XX WO200177335-A2.  
XX  
XX 18-OCT-2001.  
XX  
XX 11-APR-2001; 2001WO-FR01118.  
XX  
XX 11-APR-2000; 2000FR-0004629.  
XX  
XX (INSP ) INST PASTEUR.  
XX  
XX Buchrieser C, Frangoul L, Couve E, Rusnlok C, Faïhi H, Dehoux P;  
PI Dusurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;  
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng B, Vazquez-Boland JA;  
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;  
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;  
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;  
PI Meduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;  
PI Rose M, Voss H;  
XX  
XX WPI; 2002-010914/01.  
XX  
XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment  
PT and prevention of Listeria and related bacterial infections, and  
PT related polypeptides -  
XX  
XX  
XX Claim 6; SEQ ID No 2496; 192pp; French.  
XX  
XX The present invention relates to the genome sequence of Listeria  
CC monocytogenes EGD-e (see ABR403041). The genome sequence and fragments of  
CC it are useful for selecting probes and primers for detecting genes in L.  
CC monocytogenes and related organisms, and for studying genetic  
CC polymorphisms and other genomes. The present sequence is a protein  
CC encoded by the genome sequence of the present invention. Proteins  
CC expressed from the genome sequence are useful for raising specific  
CC antibodies, identification of L. monocytogenes and related organisms, and  
CC for biosynthesis and biodegradation, especially biosynthesis of vitamin  
CC B12. The genome sequence and proteins encoded by it are also useful for  
CC selecting compounds that regulate gene expression and cell replication  
CC and modulate L. monocytogenes-related diseases. In addition, the genome  
CC sequence and proteins encoded by it are useful in pharmaceutical and  
CC vaccines compositions for the treatment or prevention of infections by L.  
CC monocytogenes and related organisms.  
XX  
XX Note: The sequence data for this patent did not form part of the printed



PT infections and gastroduodenal diseases  
XX  
PS Claim 9; Page 90-93; 184pp; English.  
XX

CC This is the amino acid sequence of a 76 kDa Helicobacter pylori  
CC polypeptide designated GHPO 1516. It was deduced from an isolated  
CC genomic DNA sequence (see AAV07913). The invention provides a family  
CC of 76 kDa Helicobacter polypeptides (see AAV73022-32), as well as a  
CC 32 kDa polypeptide (see AAV73034) and a 50 kDa polypeptide (see  
CC AAV73035), and also polynucleotides (see AAV72001, AAV07912-21 and  
CC AAV07963-64) encoding them, expression cassettes, and methods for  
CC producing the unprocessed or mature polypeptides in host cells. The  
CC polypeptides can be used in vaccination methods to prevent or treat Hb  
CC infection in a mammal. Methods and products of the invention allow  
CC treatment and prevention of gastroduodenal diseases associated with Hb  
CC infections, including acute, chronic, and atrophic gastritis, and  
CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. Detection  
CC and diagnostic methods are also provided. GHPO 1516 was  
CC demonstrated to be a protective antigen.

Sequence 745 AA;

Query Match 5.6%; Score 100.5; DB 19; Length 745;  
Best Local Similarity 21.1%; Pred. No. 8.8; Mismatches 11;  
Matches 63; Conservative 43; Indels 81; Gaps 11;

QY 103 SASVNGSVYGFPOYLCSNFLSPNG----- 128  
DB 212 TADINGGVYQCK--AKNGSSSSSGNGSSQTATTGDDGVITTTYNNAKATVKFDI 269  
QY 129 TQASSLLELQKQVGEIV---YPDVASSSFVFGLYQQLLOSSSAVDIKASDLPQ 185  
DB 270 TNNABQLNQANI--MVLNTQCPVRSNTNMENTPGGQPGWGLSTSGNACSI---FOQ 323  
QY 186 SGDOVNRKDIQKVRITLSTVASOREYNSVKGKPISNYY-VGSESMCEIKDIIIRDQ 244  
DB 324 EFSOVTSMIKNAQETIAQSKIVSENAQNNLDGKFPNPTDASFAQSM-----LKNA 377  
QY 245 QYNVQLIGTSDKPYVYTVLALNSNLCDL-----KQYAVEVYIKNL----- 286  
DB 378 QAQAEMLPFLSEQVKNLEVMKNNNNVNEKLAGFGKEVMTNVSFAFLASCKDGTLPNAG 437  
QY 287 --TNTLVLDLGLGLTPANKGIAHLAKSSNFYAQLSQQPDASEVRLRCVDFANK 343  
DB 438 VTSNTWAGACAYVGETTISALNTSIH-----FGTQEQIQQAENIADTL--VNFKSR 487

Search completed: January 24, 2003, 19:41:47  
Time : 51 secs